UNIVERSITÄT BERN

Graduate School for Cellular and Biomedical Sciences

University of Bern

Genetic analysis of genodermatoses

PhD Thesis submitted by

Anina Estrella Bauer

from **Zürich, ZH**

for the degree of

PhD in Biomedical Sciences

Supervisor

Prof. Dr. Tosso Leeb

Institute of Genetics

Vetsuisse Faculty of the University of Bern

Co-advisor

Dr. Arnaud Galichet

Department for BioMedical Research

Faculty of Medicine of the University of Bern

Accepted by the Faculty of Medicine, the Faculty of Science and the Vetsuisse Faculty of the University of Bern at the request of the Graduate School for Cellular and Biomedical Sciences

Bern, Dean of the Faculty of Medicine

Bern, Dean of the Faculty of Science

Bern, Dean of the Vetsuisse Faculty Bern

Abstract

The skin is the largest organ of the body. It provides a barrier that protects the body from harmful environmental factors as well as from a loss of fluids. Impaired skin integrity can lead to a disruption of the skin barrier and to disease. Genodermatoses are a heterogenous group of mostly rare single gene disorders affecting the skin and its appendages. After decades of genetic and due to recent advances in technology, the genetic cause of numerous genodermatoses was uncovered. Based on the function of the involved genes and the affected skin compartments, monogenic skin disorders can be grouped into ichthyoses, epidermolysis and blistering disorders, pigmentation disorders, disorders of ectodermal appendages, vascular disorders, connective tissue defects, dermal mosaics, and genodermatoses with tumor predisposition. Many human genodermatoses have a very similar counterpart in animals. Spontaneous mutants in purebred animals such as dogs that have a unique population structure and physical similarity to humans, are therefore valuable models for human genodermatoses.

In this thesis, I took part in the analysis of 12 canine, equine and feline phenotypes with manifestations in skin or its appendages applying different genetic mapping techniques and whole genome sequencing. In cats, a frameshift variant in the *COL5A1* gene was identified in a single case with Ehlers-Danlos syndrome. In a female cat with inflammatory linear epidermal lesions, a missense variant in the X-chromosomal *NSDHL* gene was found, explaining the observed cutaneous mosaicism. Interestingly, I identified a large deletion in the same gene, *NSDHL*, underlying a related congenital cornification disorder in a Labrador Retriever and her equally affected crossbred daughter. In a single canine case of ichthyosis, whole genome sequencing revealed a *de novo* variant in the *ASPRV1* gene, probably disturbing filaggrin processing during cornification. A deregulation in cornification of nasal keratinocytes was also suspected in a litter of Greyhounds with nasal parakeratosis, in which I identified a splice defect in the *SUV39H2* gene. Pigmentation and hair texture are diverse in dogs. During a study in Chow Chows with coat colour dilution I identified a variant in the *MLPH* gene encoding melanophilin, and found that this variant segregated in other dog breeds as well. In dogs with oculocutaneous albinism, variants in *OCA2* and *SLC45A2* encoding transporters in membranes of melanosomes were identified. In dogs with a previously unknown cause for their curly coat, a candidate gene approach led to the identification of a second *KRT71* allele for curls, which has a potential predisposing role for hair loss that warrants further investigation. Hairless skin patches on the head and back was also a prominent feature in male Dachshund puppies with X-linked hypohidrotic ectodermal dysplasia, where a single base deletion in the *EDA* gene was found using targeted Sanger sequencing. Lethal

acrodermatitis in Bull terriers is a disorder causing early death of affected animals. All cases in our cohort were homozygous for a splice defect in *MKLN1*, a gene not yet described in human genodermatoses. Finally, in horses from the Akhal-Teke breed, I identified a nonsense variant in the *ST14* gene most likely causing the lethal naked foal syndrome in a monogenic autosomal recessive mode.

The identification of these candidate causative variants enables genetic testing, controlled breeding and in the long term eradication of the corresponding disorders from the animal population. The majority of identified variants was located in genes already known from human disorders. However, *ASPRV1* and *MKLN1* variants have never been reported as cause for human genodermatoses. This thesis therefore demonstrates that genetic anlysis of spontaneous animal mutants offers the chance to gain new biological knowledge and candidate genes for rare human genodermatoses.

Table of Contents

Introduction

Monogenic skin disorders, also called genodermatoses, are a group of rare single gene disorders. The skin, the organ that is by definition affected in these disorders, is briefly explained in the first background section of this thesis. This section is based on human physiology. In the second introductory part, a more detailed discussion of the skin compartments affected in different groups of skin disorders is given. As the topic of this thesis is very broad, I will focus on background information that turned out to be relevant for the thesis. For each group of genodermatoses, only a few selected examples of human or animal disorders are given. A more extensive compilation of human and animal genodermatoses can be found in the appendix. In the last section of the introduction, I summarize the basics and most relevant aspects of different methods and strategies for the identification of candidate causative variants.

Skin structure and function

The skin is the largest organ of the body. It fulfills four key functions: First, it protects the body from harmful external factors such as invading microorganisms. Second, it provides an inward barrier and prevents the loss of water and other molecules. Third, it regulates the body temperature and, fourth, allows sensory perception. The structure and integrity of the skin are critical to maintain these functions. As shown in Figure 1, the skin can be divided into three layers: the epidermis, the dermis and the subcutaneous tissue (1).

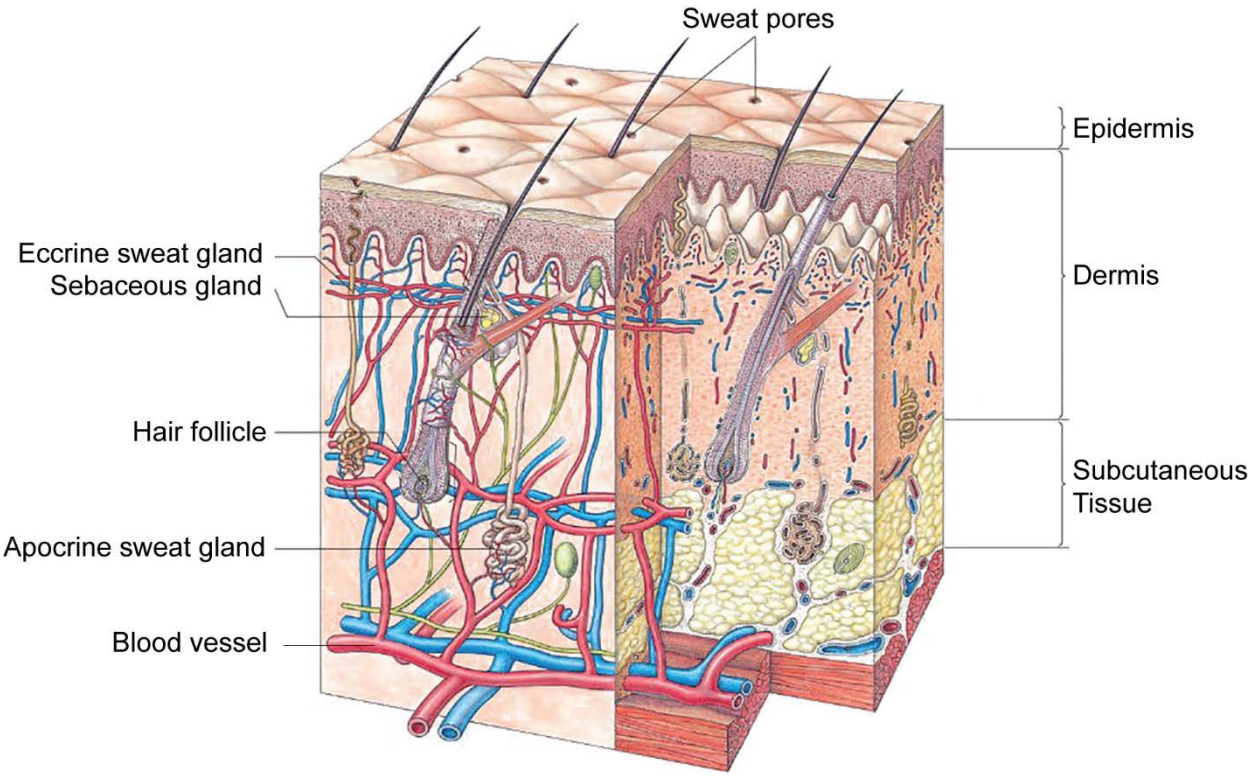
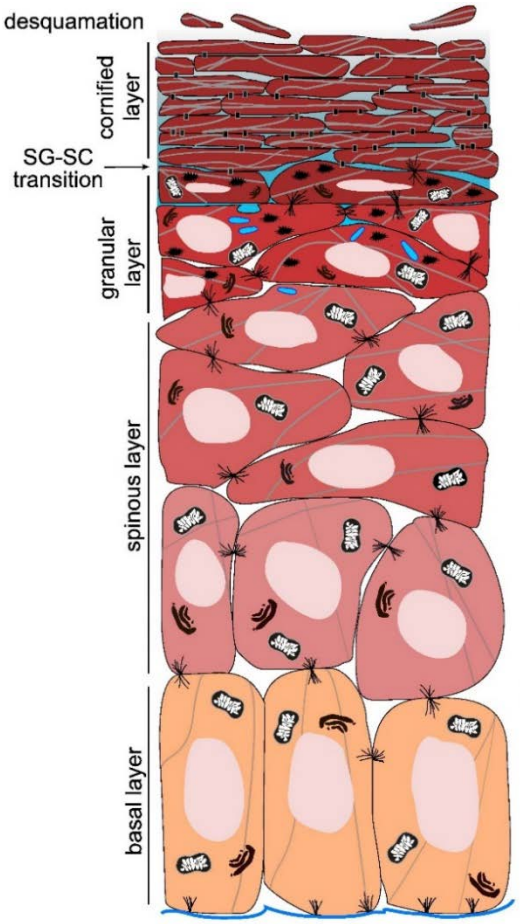


Figure 1: Structure of human haired skin. Reproduced with modifications from (1). The three layers of the skin are the epidermis, dermis and subcutaneous tissue.

Epidermis: In the epidermis, about 95% of the cells are keratinocytes. Other cell types present are melanocytes that are required for melanin synthesis, antigen presenting immune cells called Langerhans cells, and Merkel cells, involved in touch sensation. The epidermis is a stratified epithelium that undergoes self-renewal throughout life. In the different epidermal layers shown in Figure 2, the keratinocytes are present at distinct differentiation stages. On the bottom, the basal layer or stratum basale consists of proliferating keratinocytes and the



epidermal stem cells. Basal cells are anchored to the basal membrane. The layer above the basal layer is the spinous layer or stratum spinosum, followed by the granular layer, also called stratum granulosum. Typical for the granular layer are the keratohyalin granules that contain mainly profilaggrin (discussed later), as well as lamellar bodies that are lipid-rich. The uppermost layer of the epidermis, the cornified layer or stratum corneum, consists of flat keratinocytes undergoing terminal differentiation into dead corneocytes (1). In some parts of the body with thick skin, such as the palms and soles, an additional layer, the stratum lucidum or clear layer is present between granular and cornified layer. Proliferation, differentiation and shedding of corneocytes are tightly regulated and must be in balance for maintenance of epidermal function and homeostasis.

Figure 2: Layers of the epidermis. Adapted from (2).

Dermis: The dermis is the layer below the epidermis. The cellular components of the dermis are fibroblasts, macrophages, mast cells, plasma cells, endothelial cells and nerve cells. The major part of the dermis is however the dermal matrix, composed of collagen fibers, elastic fibers, reticular fibers and a gelatinous matrix. The dermal matrix is mainly produced by fibroblasts. The dermis can be divided into the papillary layer, the subpapillary layer and the reticular layer (Figure 3). The papillary layer is the uppermost part of the dermis that projects into the epidermis and allows an exchange with the epidermis. The papillary layer and the underlying subpapillary layer are rich in capillary blood vessels, sensory nerve endings and fibers. The main part of the dermis is part of the reticular layer. It is located between the subpapillary layer and the subcutaneous tissue. In this layer, the connective tissue is very dense, consisting of different fibers (1).

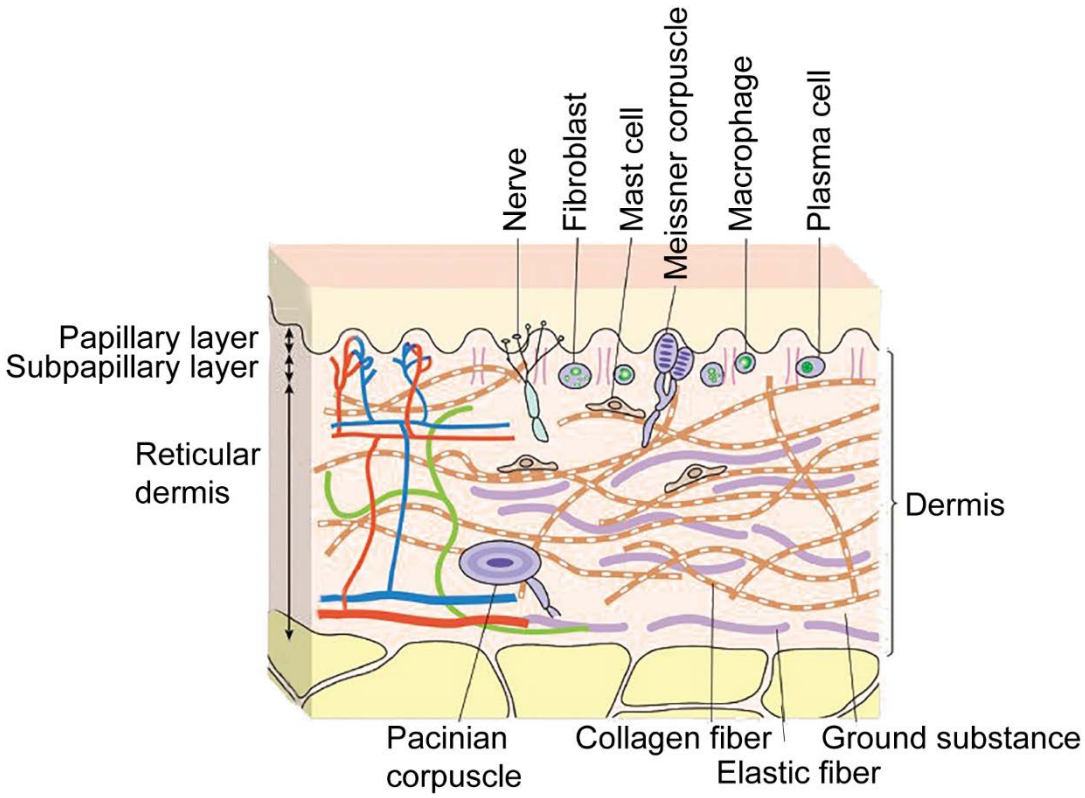


Figure 3: Layers and main components of the dermis. Reproduced with modifications from (1). The dermis is divided into papillary, subpapillary and reticular layers. Collagen, elastic fibers and the ground substance are part of the dermal matrix. Cellular components shown are fibroblast, mast cells, macrophages and plasma cells. Pacinian and Meissner corpuscles are sensory nerve terminals (1).

Subcutaneous fat tissue: The subcutaneous tissue is located under the dermis and is mainly composed of fat cells. It functions as a cushion against external physical pressure and is important for thermo-regulation. The presence or absence and the thickness of the subcutaneous tissue is age- and site specific (1).

Skin appendages: Skin appendages are ectoderm-derived organs that differ from each other in function and shape but share similar developmental pathways. In humans, these ectodermal derivatives include hair, teeth, nails, and eccrine glands such as sweat, salivary or mammary glands. These ectodermal organs typically have the potential to regenerate (3). Like the skin, appendages differ between species. As an example, humans have mainly hair follicles with a single hair shaft, while for example in dogs compound hair follicles predominate (4).

Monogenic skin disorders

Monogenic skin disorders, also known as genodermatoses, are single gene disorders with cutaneous manifestation. The Online Mendelian Inheritance in Man (OMIM) database (https://www.omim.org) listed 560 genodermatoses with known genetic etiology in 2009 (5). This number has been increasing during the last decade, mainly due to improvement in technologies such as next generation sequencing (NGS) (6). More than 80% of monogenic skin disorders are caused by genetic defects in only one known gene, and defects in approximately 80% of identified genes in genodermatoses cause only one disorder. The majority of genodermatoses has systemic manifestations (5). In 2014, Lemke and colleagues published an overview of known monogenic human skin disorders and their underlying genetic cause (6). Based on the phenotypic spectra of the disorders, nine groups of genodermatoses were defined, which will be discussed in the following sections.

Inherited ichthyoses / generalized Mendelian disorders of cornification

In the epidermis, keratinocytes move from the basal layer up to the cornified layer, as they differentiate. The terminal differentiation, during which keratinocytes become flat, dead corneocytes without nucleus and organelles, is also known as cornification or keratinization (7). The main purpose of the cornified layer is to function as a physical and permeability barrier (8). This barrier is formed by corneocytes that are tightly attached to each other and the surrounding lipid rich extracellular matrix. Eventually, the dead cells are shed from the skin surface in a process called desquamation. Inherited ichthyoses or generalized Mendelian disorders of cornification (MeDOCs) are genodermatoses caused by defects in terminal differentiation of keratinocytes and skin barrier formation (9). This category of genodermatoses comprises a heterogeneous group of disorders mainly characterized by thickening of the skin, called hyperkeratosis, and/or visible scaling. In 2014, Lemke and colleges listed 69 human disorders caused by variants in 68 different genes in this category (6).

Based on the consensus nomenclature from 2010 (10), inherited ichthyoses are often grouped into syndromic forms (11) and non-syndromic forms (12). I will however briefly discuss the major components of the epidermal barrier and give examples for skin disorders resulting from genetic variants in genes required for their formation. As shown in Figure 4, these components are the intracellular keratin network, the cornified cell envelope and the intercellular lipid layer (13, 14). Important, and to add as a fourth component, are the structures called corneodesmosomes responsible for the adhesion of the corneocytes, as well as their proteolytic degradation to allow shedding of the corneocytes (15, 16).

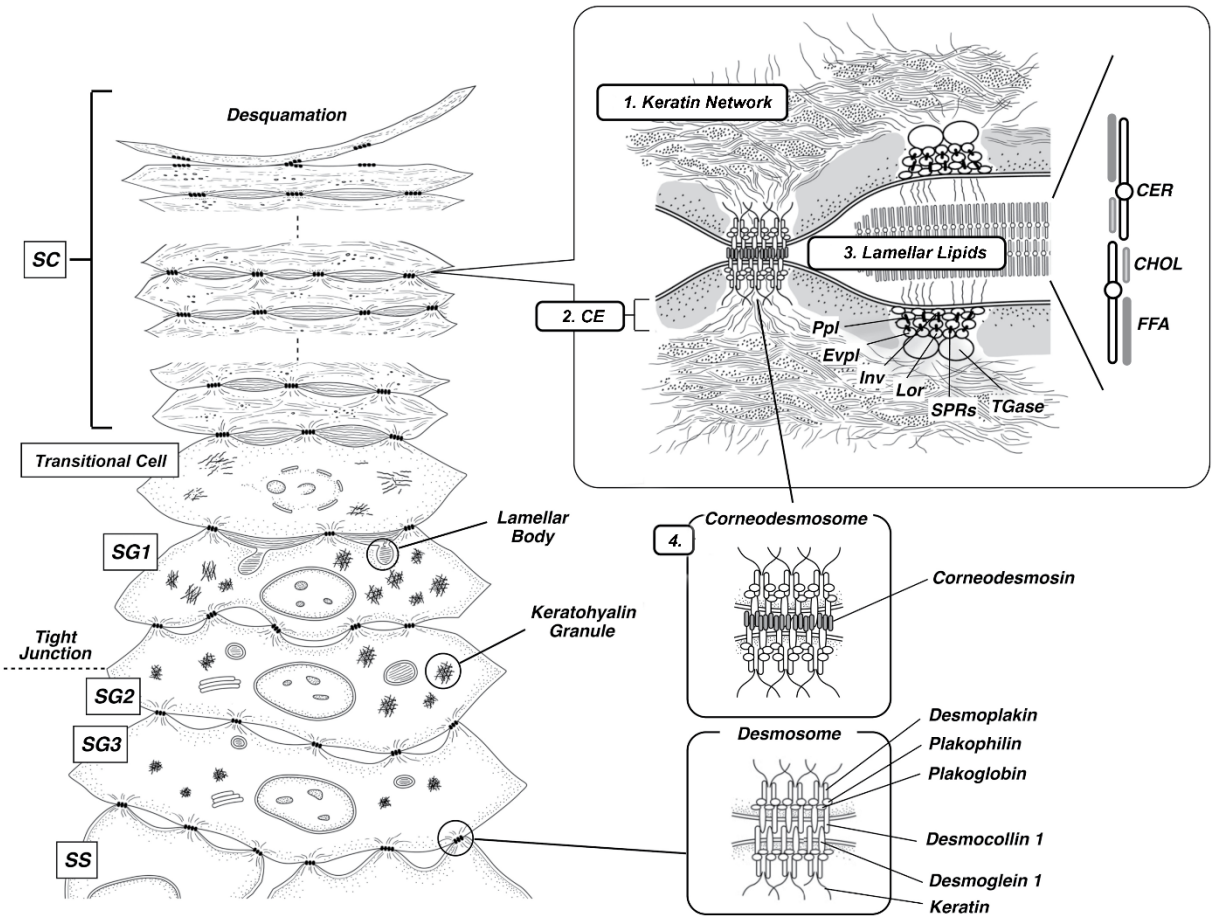


Figure 4: **An illustration of the major events taking place during cornification. Reproduced with modifications from (13).** 1. Formation of the keratin network, 2. cornified envelope formation, 3. formation of intercellular lipid layers (cornified lipid envelope not shown), 4. transformation of desmosomes into corneodesmosomes. Keratohyalin granules and Lamellar bodies are shown in in keratinocytes of the granular layers. CE= cornified envelope, CER = ceramide, CHOL = cholesterol, Evpl = envoplakin, FFA = free fatty acid, Lor = loricrin, Ppl = periplakin, SS = stratum spinosum, SG = stratum granulosum, SC = stratum corneum, TGase = transglutaminase.

The intracellular keratin network: As keratinocytes differentiate, they express different keratin intermediate filament proteins. In the basal layer, mainly keratin 5 and 14 are expressed. In the suprabasal layers, they are replaced with keratin 1 and 10 early during cornification (7). Keratins are expressed in a stage- but also tissue-dependent manner. As an example, keratin 9 is specific to palms and soles where it provides additional protection against mechanical stress (17, 18). Pathogenic genetic variants in the *KRT9* gene therefore typically cause an autosomal dominant skin disorder called epidermolytic palmoplantar keratoderma (OMIM #144200), characterized by diffuse thickening of the palmoplantar epidermis and erythema (19). The generalized form of the disorder, epidermolytic hyperkeratosis (OMIM #113800) has been reported in patients with *KRT1* and *KRT10* variants, present in either a heterozygous or a homozygous state (20-24). However, *KRT1* variants can also cause palmoplantar keratoderma, similar to the phenotype observed in individuals with *KRT9* variants.

At a later stage in differentiation, keratinocytes acquire keratohyalin granules. These granules contain mainly profilaggrin, a phosphorylated large precursor polyprotein that will be processed into functional filaggrin (filament-aggregating protein) units. Profilaggrin-to-filaggrin processing requires dephosphorylation and proteolysis by multiple enzymes (25). Filaggrin is further processed into free amino acids, components of so called natural moisturizing factors (26). The typical flattened shape of corneocytes is due to the tight bundling of keratin intermediate filaments by filaggrin. Together, filaggrin and keratin contribute approximately 80-90% of total protein mass of the mammalian epidermis (7, 27).

The most common inherited ichthyosis is ichthyosis vulgaris (OMIM #146700) caused by variants in the *FLG* gene encoding profilaggrin. Ichthyosis vulgaris is one of the most common monogenic human disorders and is inherited in an autosomal semidominant mode with incomplete penetrance. The characteristics of the relatively mild disorder are palmar hyperlinearity, follicular hyperkeratosis and no to prominent scaling (28). As not only filaggrin itself but also the profilaggrin processing steps are critical during cornification, genetic defects in profilaggrin- or filaggrin- processing enzymes can also lead to impaired barrier function (29).

Cornified envelope formation: The cornified envelope is an approximately 10 nm thick insoluble layer formed beneath the plasma membrane. It consists mainly of cross-linked proteins. The isopeptide bonds cross-linking the proteins are the main reason for the insolubility of the cornified envelope, and they are formed by transglutaminases (TGMs) (30, 31). In suggested models for envelope formation, the proteins involucrin, envoplakin and periplakin are expressed and cross-linked earlier in differentiation, followed by loricrin, elafin, sPRs and S100 proteins as well as the late envelope proteins (30-33). Expression and cross-linking of the different envelope proteins by TGMs is influenced by the raising Ca^{2+} concentration during differentiation of the keratinocytes. Among the nine known human TGMs, three of them have been shown to be involved in cornified envelope assembly, namely TGM1, TGM3 and TGM5. Variants in the *TGM1* gene cause autosomal recessive lamellar ichthyosis (OMIM #242300) in humans (34, 35). Characteristic for this disorder is the so-called collodion membrane at birth, presenting with plate-like, large, brown scales (7). *Tgm1* knockout mice die within 4-5 hours after birth due to dehydration (36). This demonstrates that, in specific tissues, some TGMs have functions which cannot be compensated by other TGMs. A defect related to structural proteins of the cornified envelope is e.g. a variant form of Vohwinkel syndrome (OMIM #604117) caused by dominant variants in the *LOR* gene encoding loricrin. This disorder is characterized by palmoplantar hyperkeratosis and constricting bands on digits of hands and feet that can lead to auto amputation (37). To date, no skin disease caused by the complete absence of a cornified envelope structural protein has been described, possibly because such traits might be embryonic lethal (7).

Corneocyte lipid envelope and intercellular lipid layer: In the cornified layer, lipids are present as a cornified lipid envelope and as intercellular lipids. The cornified lipid envelope is a single layer of ultra-long-chain ceramides and fatty acids. This layer, which is bound to the cornified protein envelope, seems to be particularly important for the barrier function of the stratum corneum. The intercellular lipids are arranged in bilayers, mainly built of ceramide, free fatty acids and cholesterols (38). Synthesis of stratum corneum lipids begins in the endoplasmatic reticulum of stratum spinosum keratinocytes. Newly synthesized lipids are modified in the Golgi apparatus and packed into lamellar bodies, Golgi-derived vesicles. In the upper granular layer, the vesicles release the lipids and other contents such as hydrolases involved in the extrusion process, into the extracellular space (7, 39). Defective lipid layers can be due to defects in synthesis, metabolism, transport, secretion or arrangement of epidermal lipids. In 2012, Grall et al. identified an autosomal recessive defect in the *PNPLA1* gene in Golden Retrievers with ichthyosis and subsequently in humans with autosomal recessive congenital ichthyoses (ARCI10; OMIM #615024) (40). *PNPLA1* encodes patatin like phospholipase domain containing protein 1, a transacylase required for the final synthesis step of acylceramid, a unique lipid in the stratum corneum barrier (41).

Corneodesmosomes and desquamation: During cornification, desomosomes are transformed into corneodesmosomes, which are the main adhesive structures in the cornified layer of the epidermis. Corneodesmosoms are structurally but also compositionally distinct from desmosomes. The main difference in composition between desmosomes and corneodesmosomes is that corneodesmosin is present in the latter (15). It has been shown that if this component is missing, the cornified layed detaches prematurely from the granular layer in mice, and that they die within a few hours after birth (42). Homozygosity for loss of function variants in the *CDSN* gene encoding corneodesmosin is not lethal in humans, but leads to peeling skin syndrome 1 (OMIM #270300) as first described by Oji et al (10). *CDSN* variants in a heterozygous state can lead to hypotrichosis of the scalp (OMIM 146520), consistent with the expression of corneodesmosin in the inner root sheath of hair follicles (43). To maintain a certain thickness of the epidermis, proliferation in the basal layer has to be balanced by desquamation in the uppermost layer. Desquamation occurs by complete degradation of the extracellular corneodesmosome components corneodesmosin, desmoglein 1 and desmocollin (15). Important molecules in this process are the serine proteases kallikreins (KLKs) and the lympho-epithelial Kazal-type related inhibitor 1, LEKTI, encoded by *SPINK5* (44, 45). Netherton syndrome (OMIM #256500), a severe autosomal recessive disorder, is caused by variants in *SPINK5* (46). Characteristics for this syndrome are congenital erythroderma, sparse and brittle scalp hair and atopic manifestations with high IgE levels (47).

Inherited epidermolyses and blistering disorders

Intact cell-cell and cell-matrix attachments are important for the integrity and mechanical strength of the skin and other stratified or complex epithelia. Structures responsible for cell-cell adhesion include desmosomes, adherens junctions, gap junctions and tight junctions (48). At the dermal-epidermal junction or basement membrane zone, epithelial cells are attached to the basement membrane, which is attached to the papillary dermal layer. Hemidesmosomes and focal adhesions are structures that are crucial for this attachment but also play a role in different signaling pathways. A genetic defect in a gene encoding a molecule required to build these multiprotein structures can lead to an impaired mechanical strength of the skin, which typically leads to skin fragility, blistering and/or inflammation after minor mechanical trauma. On the other hand, blistering diseases can also be acquired and caused by autoantibodies as in pemphigoid diseases (49).

The largest group of blistering diseases is termed Epidermolysis Bullosa (EB). According to recommendations on EB classification, there are four major types of EB: EB simplex, junctional EB, dystrophic EB and Kindler syndrome. As shown in Figure 5, this classification is based on the location of blister formation in the skin: Intraepidermal in EB simplex, within or beneath the basement membrane zone in junctional EB and dystrophic EB, respectively, or with a mixed pattern in Kindler syndrome.

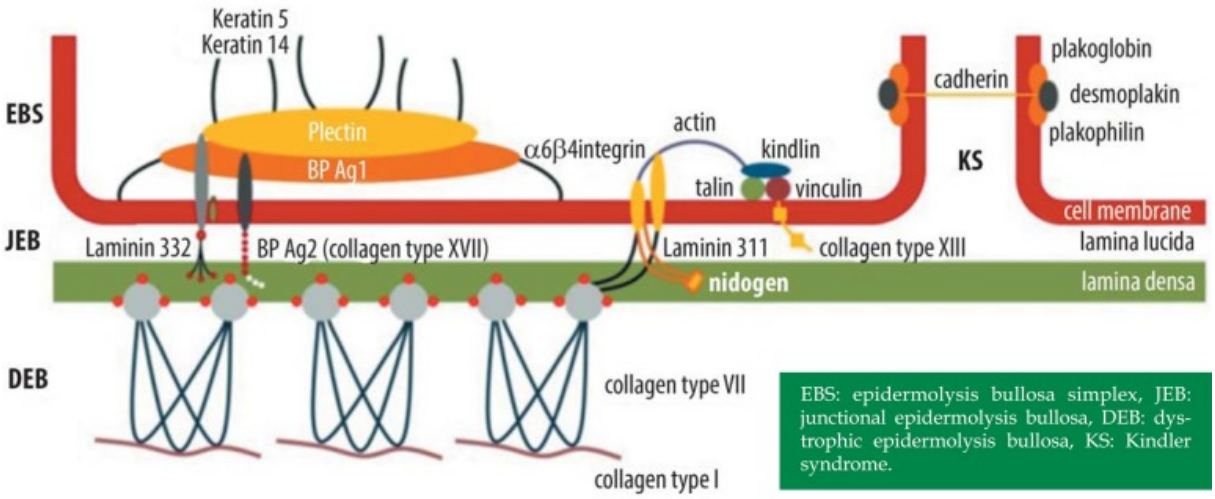


Figure 5: illustration of adhesion structures at the dermal-epidermal junction and the proteins involved in different forms of epidermolysis bullosa. Reproduced from (50). In Epidermolysis bullosa symplex (EBS), the blister formation occurs intraepidermal,in junctional epidermolysis bullosa (JEB)genetic defects lead to a separation of the lamina lucida, and dystrophic epidermolysis bullosa (DEB) is due to defective anchoring fibrils resulting in separation of the sub-basal lamina. A mixed pattern is observed in Kindler syndrome (KS).

EB can further be classified into subtypes by mode of inheritance, clinical picture and the gene involved in the disease (51). As an example, genetic variants in *ITGA6* can cause a subtype of junctional EB (OMIM #226730) in an autosomal recessive mode of inheritance. *ITGA6* encodes the integrin α6 unit that forms heterodimers with the β4 or β1 subunits. In hemidesmosomes of the skin, this transmembrane polypeptide provides an anchorage to the

basement membrane by interacting with laminin 332. A defective integrin α6 unit therefore leads to a defect in the attachment of the hemidesmosomal plaque to the extracellular matrix and, as a consequence, to blister formation in the dermal-epidermal junction. Like other hemidesmosomal components, integrin α6β4 heterodimers are not restricted to the skin, explaining additional symptoms beyond the skin (49).

Nucleotide excision repair disorders

While the epidermal barrier offers a first protection against many exogenous agents, some factors such as UV radiation, but also endogenous processes, induce damage that requires additional repair mechanisms. Exposure to UV from sunlight causes direct DNA damage in the form of cyclobutane pyrimidine dimers and pyrimidine-pyrimidone (6-4) photoproducts. The repair mechanism that is most important in removing such products is the nucleotide excision repair (NER). As shown in Figure 6, NER has four main steps: first, the DNA damage is recognized either through the transcription-coupled (TC) or the global genome (GG) repair pathway, second, the double-stranded DNA is opened around the lesion, third, endonucleases excise the DNA lesion, and fourth, new nucleotides are inserted by taking the complementary strand as a template.

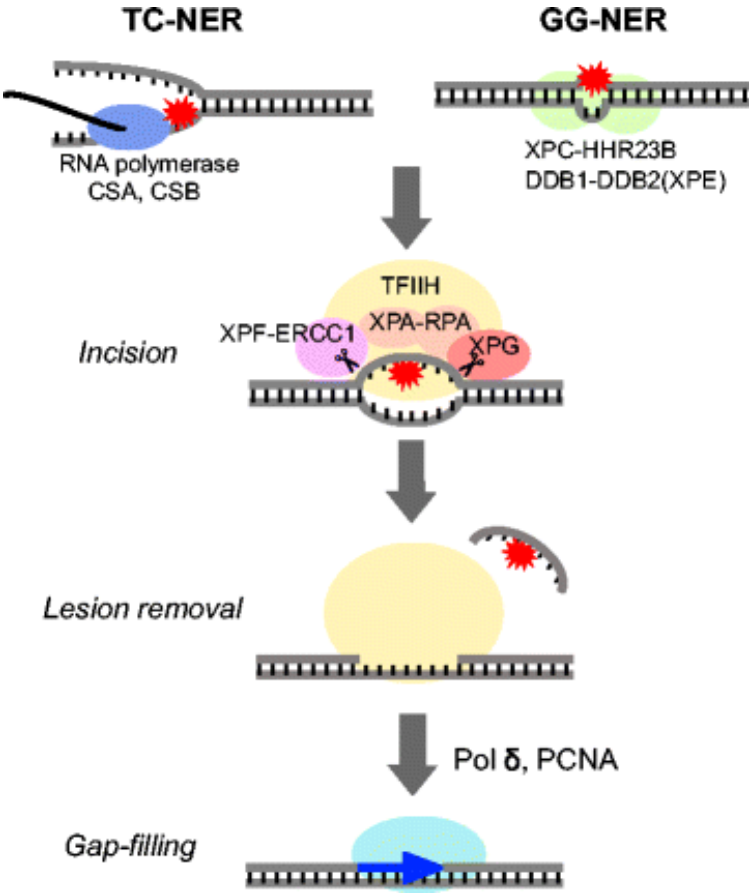


Figure 6: Simplified illustration of the main steps in nucleotide excision repair. Reproduced from (52). In reality, approximately 30 different factors are involved. The two pathways of this repair mechanism are the transcription-coupled repair (TC-NER) and the global genome repair (GG-NER). TC-NER mainly recognizes damage on the template and transcribed strands, while GG-NER repairs lesions independently of transcription (53).

NER involves approximately 30 different factors, and genetic variants in about half of them are known to cause human genetic disorders. (53). Among others, these disorders include the different subgroups of xeroderma pigmentosum, characterized by sunlight-induced changes in skin pigmentation, parchment-like dry skin, and often strong sensitivity to sunlight and photophobia (54). As an example, xeroderma pigmentosum group E (OMIM #278740) is caused by variants in the *DDB2* gene, which encodes the small subunit of the DDB complex involved in the first step of NER, the damage recognition (55, 56).

The repair of pyrimidine-pyrimidone (6-4) photoproducts happens within approximately 3 h after exposure, however, a big proportion of cyclobutane pyrimidine dimers is still present 24 h after exposure (57). To cope with this remaining damage, other repair systems such as the post replication system, also known as translesion synthesis system (TLS), exist. This mechanism is either error-prone or error-free, depending on the polymerase. In individuals with pathogenic variants in *POLH* encoding polymerase eta, the error-free axis of the TLS is not functioning normally, which leads to error-prone TLS (53, 58). This disorder, with a strong predisposition to sunlight induced skin cancer due to the increase in mutation frequency, is termed xeroderma pigmentosum, variant type (OMIM #278750) (59).

Inherited hyper- and hypopigmentation disorders

Abnormal pigmentation in skin is easily visible to the human eye, which made genetic pigmentation disorders some of the first studied traits in humans. Pigmentation requires the functioning of a complex multi-step process including melanoblast development, melanoblast migration to the skin, melanin synthesis in melanocytes, melanosome formation, and the transfer of melanosomes to keratinocytes (60). In each of these steps, multiple genes are involved and variants in many of them have been shown to cause monogenic disorders or traits in humans but also animals. Simplified and in general, genetic defects affecting development, migration and survival of melanocyte precursors lead to a complete or partial absence of melanocytes in skin and other target sites, resulting in disorders with abnormal pigmentation and additional clinical signs such as deafness. Defects in later steps in the process of pigmentation such as melanin synthesis or melanosome trafficking typically have an influence on pigmentation only, mostly causing non-syndromic phenotypes.

Melanoblast development and migration: The pigment melanin is the main factor influencing skin and hair colour in mammals. It is produced by melanocytes, which develop from their neural crest cell derived precursors, the melanoblasts. The melanoblasts migrate to their target sites, the epidermis and hair follicles as well as some parts of the inner ear and the eye,

where they proliferate and differentiate into melanin producing melanocytes (60, 61). Different signaling pathways and transcription factors are involved in the development, migration and differentiation of melanoblasts. Key players include the *PAX3, SOX10, MITF, KIT, EDN3* and *EDNRB* genes (60). Among those, *MITF*, encoding melanogenesis associated transcription factor with a basic helix-loop-helix leucine zipper motif, is often referred to as the key regulator gene (62). In humans, variants in *MITF* cause Waardenburg syndrome 2A (WS2A, OMIM #193510), a disorder with pigmentary and auditory abnormalities inherited in an autosomal dominant mode (63). Other disorders caused by *MITF* defects are the similar but more severe Tietz albinism-deafness syndrome (OMIM #103500) as well as the autosomal recessively inherited COMMAD syndrome (OMIM #617306) reported in children from parents with WS2A (64, 65). Phenotypes comparable to WS2A were reported in animals, including white spotting in horses that may be accompanied by deafness (66).

Melanin synthesis: Mature melanocytes have lysosome-related organelles called melanosomes, where melanin is produced. Two types of melanin can be distinguished in skin of mammals: eumelanin, a dark black-brown pigment and pheomelanin, which is yellow-red (60). Melanin is produced by melanogenic enzymes: the tyrosinase (TYR) and tyrosinase-related proteins (TYRP1 and TYRP2). The autosomal recessive disorders caused by complete or partial loss of tyrosinase activity are oculocutaneous albinism (OCA) type 1A (OMIM #606933) and 1B (OMIM #606952), respectively. These forms of true albinism are characterized by loss or reduction of pigmentation in skin, hair and eyes. Other less severe types of OCA with some amount of pigmentation are caused by variants in *OCA2, TYRP1* and *MATP* as reviewed by Gronskov et al. (67).
Melanogenesis and the switch between eumelanin and pheomelanin are regulated by other proteins such as melanocortin 1 receptor (encoded by *MC1R*), its negative agonist agouti signaling protein (encoded by *ASIP*), and other ligands (68). As reviewed by Kaelin and Barsh, the effects and interactions of molecules involved in the synthesis and transport of eu- and pheomelanin can be studied perfectly in domestic animals such as dogs and cats, as they display a large variety of different coat colours and patterns (61). Loss of function variants in *MC1R* can cause a red/yellow coat colour consistent with a switch from eumelanin to pheomelanin, while constitutively activating variants can cause a black coat due to increased eumelanin production (69). In Figure 7, the main (canine and feline) players in melanogenesis and melanosome transport are shown.

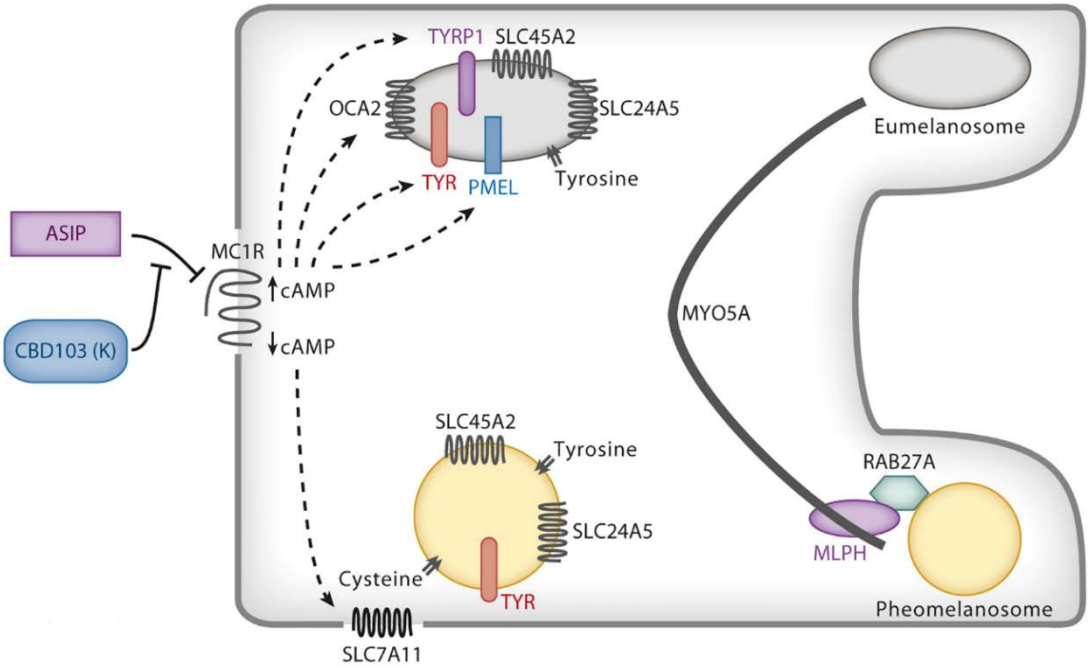


Figure 7: Simplified overview of genes involved in melanin synthesis and melanosome transportation. Reproduced from (61). Defects in many of these genes give rise to pigmentary phenotypes. *MC1R* controls the switch between eumelanin and pheomelanin production via second messenger cAMP. High *MC1R* expression and high cAMP levels increase *TYRP1*, *OCA2, TYR, PMEL* expression and lead to increased eumelanin production. Increased pheomelanin production is promoted by low cAMP levels and high expression of *SLC7A11*, encoding a cysteine transporter. *MLPH, RAB27A* and *MYO5A* are involved in melanosome trafficking.

Melanosome formation: Although melanosomes are lysosome related organelles, they are built from early endosomal intermediates. In the skin, melanosomes mature within the melanocytes. They go through four stages and move to the periphery of melanocytes before they are transported to surrounding keratinocytes. In stage I, the premelanosomes contain irregular proteinaceous fibrils that become fully organized in stage II. These intraluminal structures form a template for melanin deposition. Melanin synthesis begins in stage III and melanin accumulates, resulting in melanin-packed melanosomes in stage IV (70, 71).The machinery and components for melanosome biogenesis and melanin synthesis have to be imported into the organelle. One of these structural components is PMEL17 encoded by *PMEL* (formerly also known as *SILV*). PMEL17 is the main component of the fibrils in stage I and II melanosomes. In dogs, Clark and colleges identified that the merle coat pattern is caused by a SINE (short interspersed nuclear element) insertion in the *PMEL* gene. Merle is characterized by spots of diluted pigment and can be accompanied by ocular and auditory abnormalities, mostly but not exclusively in homozygous mutant dogs (72). In humans, a group of diseases with a defect in biogenesis or trafficking of lysosome related organelles, most importantly platelet dense granules and melanosomes, is termed Hermansky Pudlak syndrome. Clinical signs for this heterogeneous group of autosomal recessive disorders are oculocutaneous albinism, a bleeding tendency and lysosomal ceroid deposition (73).

Melanosome transfer to keratinocytes: After the development into mature, fully melanized organelles, melanosomes are transported from the center to the periphery of the melanocytes. From this location, they will be transported further to keratinocytes. The long-range transport within the melanocyte, and towards the periphery, occurs along microtubules with kinesin-2 as motor protein. Near the cell periphery, the transport continues along actin fibers with the motor protein MYO5A. The melanosome is bound to MYO5A via two adaptor proteins: melanophilin, encoded by *MLPH*, and the small GTPase RAB27A. Defects in any of the three proteins in this complex lead to abnormal pigmentation (74). How melanosomes are transferred from melanocytes to keratinocytes is still incompletely understood, and several models have been suggested (75). A recent study demonstrated that in chicken embryonic skin, melanosomes are transported in vesicles produced by plasma membrane. These vesicles are formed by blebbing of the melanocyte membrane, a process in which ras homolog family member A, encoded by *RHOA,* is involved (76). It is further known that the G-protein coupled receptor F2RL1, also known as PAR2, is involved in the phagocytosis of melanosomes in a Rho-dependent manner (77).

Disorders of ectodermal appendages

Ectodermal appendages develop through a crosstalk of epithelial and mesenchymal tissues. The early developmental stages are similar between the different appendages and species, and are mainly regulated by members of the fibroblast growth factor, hedgehog, transforming growth factor β (TGFβ), tumor necrosis factor and Wnt pathways (78, 79). A good example for a disorder affecting several ectodermal appendages that demonstrates common steps in early ectodermal appendage development is hypohidrotic ectodermal dysplasia (OMIM #305100). This disorder is caused by variants in genes of the ectodysplasin signaling pathway, most frequently in the *EDA* gene encoding ectodysplasin. Ectodermal dysplasia is known in different species such as mice, cattle, dogs and humans, and manifests with similar clinical signs including malformed or missing teeth, sparse hair, and defects in different glands leading to an inability to sweat (80-83).

For the disorders studied in the framework of this thesis, especially the hair and hair follicle are of importance. After formation of the hair follicle during embryonic development, the hair follicle goes through a lifelong cycle of growth (anagen), regression (catagen) and quiescence (telogen) as shown in Figure 8. (84). The morphology of the hair follicle changes during the phases of the hair cycle. The major compartments of a mature anagen hair follicle are the dermal papilla composed of dermal cells, epidermal matrix cells that move upwards while

differentiating into the cortex and medulla of the hair shaft as well as into the inner root sheath surrounding the hair shaft, and an outer root sheath that surrounds the inner root sheath. The basement membrane separates the epithelial cells from the dermis and the dermal papilla. The hair follicle bulge contains epithelial stem cells that have the potential to regenerate the hair follicle after the resting phase. The signals for this transition into a new anagen phase are mainly sent from dermal papilla cells, and received from the epithelial stem cells (85, 86).

Genetic hair disorders can be caused by defects in hair follicle morphogenesis, cycling, or structural components of the hair shaft or follicle (87). A classic example for a disorder due to a structural defect is monilethrix (OMIM #158000) characterized by fragile, easily braking hair resulting in scarring alopecia. This autosomal dominant disorder is caused by variants in the *KRT81, KRT83* and *KRT86* genes encoding hair keratins expressed in the mid-cortex of the hair shaft (88-91). Among the 54 human keratin genes, 17 encode hair keratins and nine encode hair follicle-specific epithelial keratins; many of them involved in hair disorders (91).

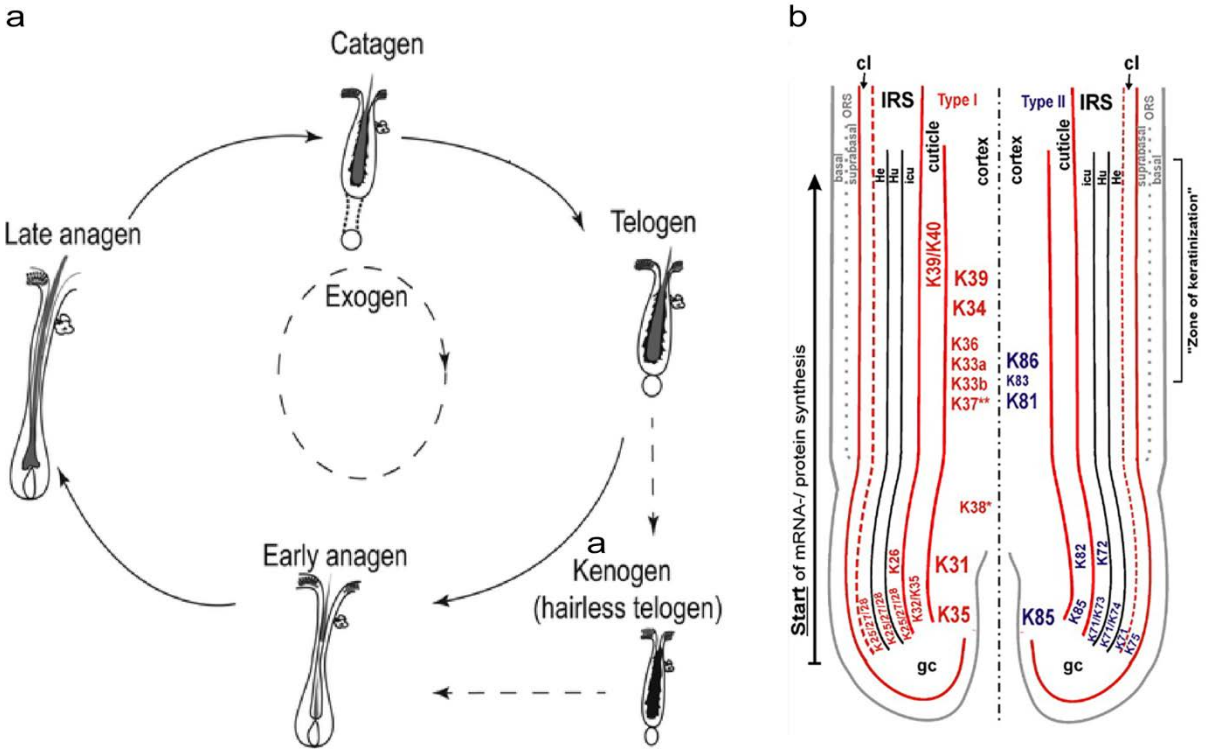


Figure 8: Hair cycle and expression of keratins in different layers of the hair follicle. Reproduced from (91, 92). a) Hair follicles cycle through phases of growth (anagen), regression (catagen) and quiescence (telogen). Kenogen refers to hair follicles that remain hairless for a certain time after telogen. b) Type 1 (left side) and type 2 (right side) keratin expression in the different layers of the hair follicle. cl = companion layer, gc = germinative center, He = Henle layer, Hu = Huxley layer, icu = inner cuticle, IRS = inner root sheath, ORS = outer root sheath.

Vascular disorders

The vessels of the blood and lymphatic system are made of a single luminal layer of endothelial cells and surrounding layers consisting of vascular smooth muscle cells and/or pericytes (93). Vasculogenesis, angiogenesis and lymphangiogenesis are the processes by which blood and lymphatic vessels are built (Figure 9). Vasculogenesis refers to the process in the embryo, where a so-called primary capillary plexus is formed de novo through differentiation of mesodermal precursor cells. During angiogenesis, this pre-existing primary capillary plexus is remodeled to form a complex network of mature capillaries, veins and arteries (94). Lymphangiogenesis is the incompletely understood process during which the lymphatic system is built. After angiogenesis, veins give rise to lymphatic endothelial cell precursors that build the first lymphatic structures, the embryonic lymph sacs. From these structures, the lymphatic vasculature is built (95). However, there are also indications for nonvenous derived lymphatic vasculature (96). Vascular malformations result from a defect in the development of the blood or lymphatic system. Named after the affected vessel type, they

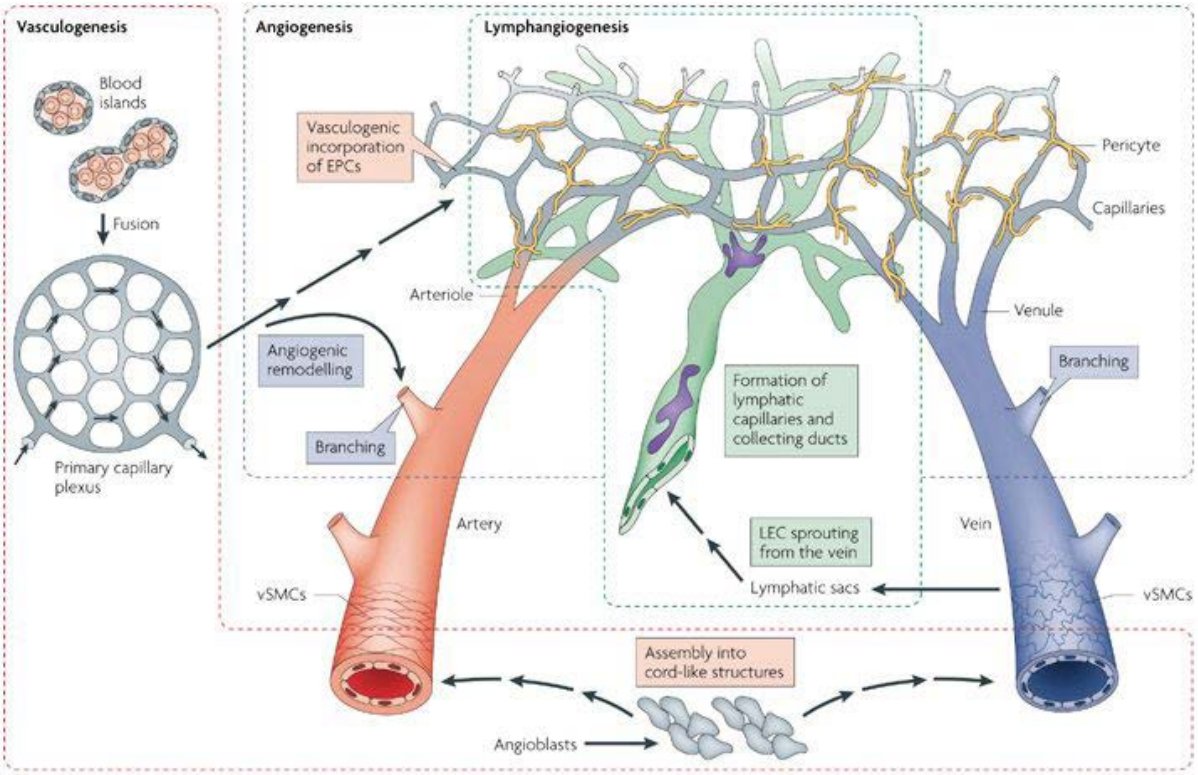


Figure 9: Events during formation of blood and lymph vasculature. Reproduced from (97). Vasculogenesis is the process of *de novo* vessel formation resulting in a primary capillary plexus. Remodeling of preexisting vessels into mature blood vessels is called angiogenesis. The formation of lymphatic vessels is called lymphangiogenesis. This process might require venous precursor cells. Pericytes are associated with capillaries; vascular smooth muscle cells cover arteries and veins. Sparse smooth muscle cells (purple) are found on larger lymphatic vessels.

can be divided into venous, capillary, lymphatic, arteriovenous and combined malformations (94). One of the most common monogenic disorders is hereditary hemorrhagic telangiectasia (HHT). This group of disorders is characterized by arteriovenous malformations in which the capillaries between veins and arteries are missing, leading to direct connections between

these two types of vessels. Compared to normal blood vessels, the abnormal vessels are typically more prone to hemorrhage because of their fragile walls that contain less vascular smooth muscle cells, and turbulent blood flow. Affected individuals often remain undiagnosed because the typical clinical signs can be subtle. First, spontaneous nosebleeds and later telangiectases, small arteriovenous malformations that are usually visible as pink to red small lesions of the skin of buccal mucosa, are the manifestations in most patients. Additional symptoms include gastrointestinal bleeding in about one third of the patients, and complications from arteriovenous malformations such as for example stroke or very rarely deep venous thromboses. HHT is an autosomal dominant disorder which is in most cases caused by variants in endoglin (*ENG*) or activin A receptor type II-like 1 (*ACVRL1*) and rarely by variants in *SMAD4, GDF2* and other unidentified genes (98-100). The products encoded by these genes play a role in the TGFβ signaling pathway. It is suggested that most variants result in haploinsufficiency and a disturbance in TGFβ signaling that is critical for angiogenesis (98, 100).

Connective tissue defects

Connective tissues function as a supportive structure in the body. They can be divided into hard connective tissues referring to bone and cartilage, blood, and soft connective tissues that surround organs. The extracellular matrix of connective tissues that embeds cellular and fibrous components consists mainly of common and tissue-specific collagen, elastin, glycoproteins and glycosaminoglycans (101). In the skin, the dermal extracellular matrix can be divided into the previously described basal membrane and the interstitial matrix (102). Based on the main type of fibers involved, connective tissue disorders are grouped into collagenopathies or elastinopathies (101). In 2014, Lemke and colleagues listed 26 genes in which variants were reported to cause monogenic connective tissue disorders related to the skin (6).

The largest group among these disorders includes different forms of Ehlers-Danlos syndrome (EDS), followed by cutis laxa, a term for a heterogenous group of elastinopathies characterized by relatively inelastic skin, loose and redundant skin folds and often multiorgan involvement (6, 103). Based on the international classification from 2017, there are 13 subtypes of EDS in humans (104). The general characteristics of EDS are skin hyperextensibility, joint hyperflexibility and skin friability (105). The classical EDS as well as the cardiac-valvular, vascular, arthrochalasia and myopathic EDS are caused by mostly but not exclusively dominant variants in different collagen genes. Most other forms are caused by variants in collagen-modifying enzymes (104).

Each collagen molecule is made of three alpha chains synthesized in the endoplasmatic reticulum as procollagen alpha chains. After different post-translational modifications, these chains associate and form a triple-helix to build a procollagen molecule. For this coiled-coil structure, at least one domain of Glycin-X-Y repeats is required in every alpha-chain. Procollagen molecules are secreted into the extracellular space where they are further processed by proteinases that cleave the N- and C-terminal propeptides resulting in tropocollagen molecules. These molecules assemble into fibrils that are stabilized by covalent cross-linking (106, 107)). Because of the assembly of collagen molecules into supramolecular structures, genetic variants resulting in structurally changed polypeptide chains that are still able to assemble with non-mutant chains usually lead to more severe phenotypes than variants, such as nonsense variants, that prevent an assembly of normal and abnormal polypeptides as illustrated in Figure 10 (106).

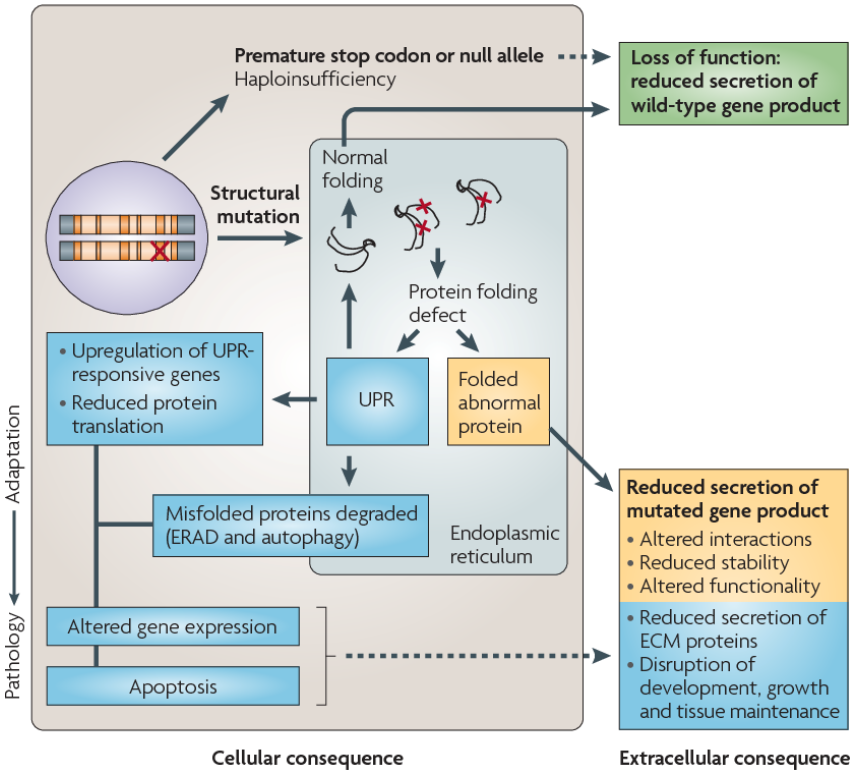


Figure 10: Molecular pathology and extracellular consequences. Reproduced from (108). Loss of function variants in genes encoding structural extracellular matrix (ECM) components such as collagen typically lead to decay of the mRNA and therefore reduced protein synthesis which may result in impaird functionality of the ECM. Variants that do not lead to a degradation of the mutant transcript but to a mutant protein that is integrated into the multimer can act in a dominant negative manner. A mutant-containing multimer can lead to protein deficiency and altered interactions, reduced stability and deleterious effects on ECM function (yellow box). If the incorrect folding leads to degradation (unfolded protein response, UPR), this can alter gene expression and lead to apoptosis and eventually disruption of tissue maintenance.

Dermal mosaics

Genetic mosaicism describes an organism with two or more genetically distinct cell populations that originated from only one zygote. This is different from chimaerism, where genetically distinct cell populations from two different zygotes come together in one individual (109). The postzygotic *de novo* mutation event leading to genetic mosaicism can occur at different developmental stages and in different body parts, leading to different inheritance patterns and disease severity. In somatic mosaicism, the disease causing variant affects only a part of the somatic cells and is therefore limited to the affected individual. If gonadal tissue is affected as well, as in gonosomal (somatic and germline) or germline mosaicism, the variant can be transmitted to the offspring (110). Some disorders are never transmitted, even if the mutation event occurred in a germline precursor cell. This can have two reasons: First, the disorder could be embryonic lethal or second, the variant could lead to apoptosis of the mutant germ cells (109). An example for such a disease that manifests only as mosaicism is the rare Proteus syndrome (OMIM #176920). Proteus syndrome is defined as a progressive, sporadically occurring overgrowth syndrome with mosaic distribution of lesions, and additional features such as for example a cerebriform connective tissue nevus (111). By whole exome sequencing of DNA originating from lesional and non-affected tissues of patients with Proteus syndrome, an activating missense variant in the *AKT1* gene was found to be present in the lesional but not the non-affected tissues. This demonstrated for the first time that Proteus syndrome is caused by *de novo* mutation events in the *AKT1* gene (112) The oncogene *AKT1* encodes AKT serine/threonine kinase 1, a central player in the PI3K/AKT/mTOR signaling pathway controlling important processes such as proliferation, cell growth and survival (113).
Besides genetic mosaicism, there is also functional or epigenetic mosaicism. One form of epigenetic mosaicism is X-chromosomal functional mosaicism, the result of early embryonic X-chromosome inactivation in females, also called lyonization. In females that are heterozygous for a X-chromosomal disease-causing variant, this mechanism will give rise to a cell population expressing the mutant allele, and cells in which the wildtype allele is expressed (114). In disorders with cutaneous manifestation, this will lead to a pattern of lesional and healthy skin patches.
In both genetic and functional mosaicism, the mosaic pattern is typically arranged along the so-called Blaschko's lines as shown in Figure 11. These lines of Blaschko might reflect the migration of skin cells during embryonic development (115).

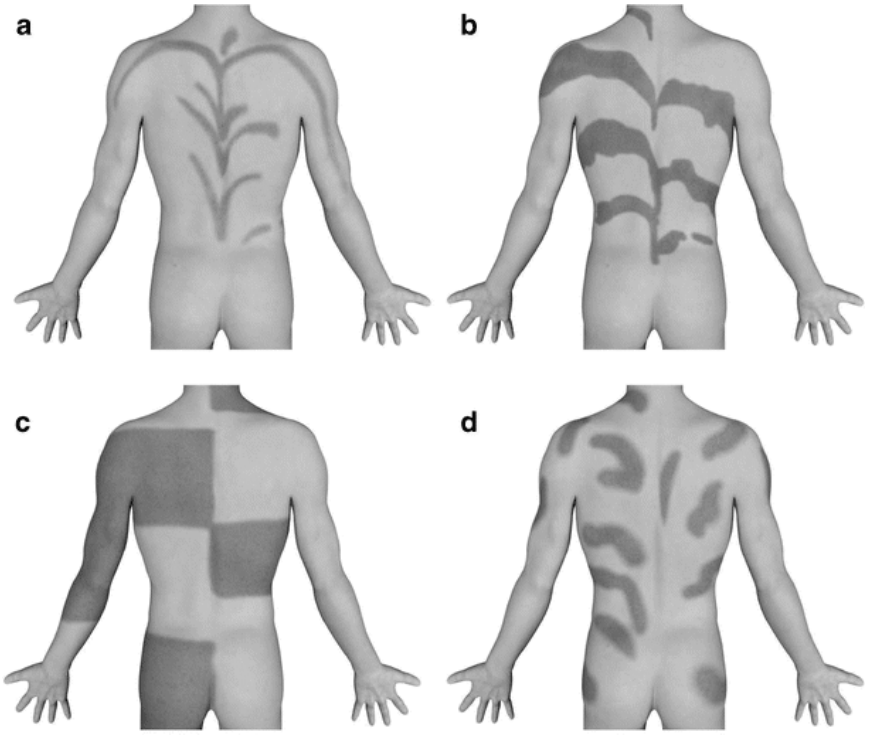


Figure 11: Different mosaic pattern following the lines of Blaschko. Reproduced from (115). Type 1: a) linear and b) broad linear lines. Type 2: checkerboard pattern or lateralization in c). Type 3: pylloid pattern in d).

Genodermatoses with tumor predisposition

Processes such as cell cycle control, DNA repair or telomere maintenance are critical for health. An imbalance caused by a mutated gene involved in such a process may lead to tumor development. On the other hand, a genetic susceptibility to carcinogenic factors such as UV light, or genodermatoses associated with immunodeficiency, might also predispose to tumors (116). Furthermore, there are genodermatoses with increased risk for cancer, such as Kindler syndrome, in which the mechanisms for cancer susceptibility are not obvious at first sight (117). Lemke et al. listed 45 genes in which variants are known to cause genodermatoses with tumor predispositions (6). Many of them are involved in the development of previously discussed xeroderma pigmentosum subtypes. Another prominent group of disorders in this list is dyskeratosis congenita. Diskeratosis congenita is a term for disorders with defective telomere maintenance, clinically characterized by mucocutaneous abnormalities (abnormal skin pigmentation, nail dystrophy, mucosal leukoplakia), increased susceptibility to cancer, and bone marrow failure, which is often the cause of death in affected individuals (118). The different subtype of the disorder can be inherited in either monogenic autosomal recessive, dominant or x-linked recessive mode (119). In all cases with a genetic diagnosis, the causative variant was in a gene coding for a component of the telomerase or telomere binding complex shelterin. As pathogenic mechanism, it is suggested that telomere dysfunction results in genome instability and tumorigenesis (120, 121).

Selected methods and technologies in genetic analyses

The first genetic causes for Mendelian disorders were identified by applying candidate gene approaches, which required biological knowledge about the disorder and its underlying etiology. In the mid-1980s, positional cloning became a popular approach to map the locus segregating with the disease. Using this approach, no biological knowledge or hypotheses were required for the initial mapping (122). Important hypothesis-free methods for the genetic analysis of Mendelian disorders include, but are not limited to, linkage analysis, homozygosity mapping, and genome wide association studies. Nowadays, these methods are based on hundreds of thousands of single nucleotide variants (SNVs) with known chromosomal positions that are used as markers. With the available genotyping arrays, genotypes for more than 1.8 million markers disributed over the whole human genome can be determined with the Affymetrix Genome-Wide Human SNP Array 6.0. In dogs, the latest genotyping array detects more than 700,000 markers. In the following paragraphs, the methods most relevant for genetic mapping of disorders investigated in this thesis will be briefly explained. In the last section, next generation sequencing will be addressed, a technology which has revolutionized the field of genetics.

Linkage analysis

Linkage analysis is a family-based method to detect chromosomal regions that are inherited together with the trait of interest. In linkage analysis, the concepts of linkage and recombination are important. Two loci in the genome are linked, if they are inherited together more often than expected by chance. This means, the probability of recombination between these loci is less than 50% during meiosis. In a genetic map, two loci separated by one centimorgan, which corresponds to approximately one billion base pairs, have a chance of 1% to be separated in one meiosis. The probability that two loci are separated during a recombination event at meiosis increases with distance between the two loci. Therefore, linkage disequilibrium between two loci decreases with every separation and only persists over several generations if the two loci are in close proximity to each other (123, 124).

Linkage analyses can be done model-free (also known as non-parametric) or parametric. Model-free linkage is based on allele-sharing and used more often for complex disorders (125). In parametric linkage analysis, the co-segregation of genetic markers and the trait of interest is investigated under a given disease model that includes information on mode of inheritance and penetrance. The strength of linkage between the trait and marker locus is usually given as LOD (logarithm of the odds) score reported by Morton in 1955 (126). Simplified, it is calculated by the logarithm to the base 10 of the likelihood to observe the data

due to true linkage (alternative hypothesis) divided by the likelihood to observe the data by chance (null hypothesis). Traditionally, a LOD score of 3 or higher is regarded as evidence for linkage, while a LOD score of -2 or lower allows to exclude linkage (123).

Homozygosity mapping / autozygosity mapping

For monogenic autosomal recessive disorders where the disease allele is expected to be identical by descent, homozygosity mapping is a powerful tool to locate the disease allele. This method is based on the assumption that an affected individual has inherited two identical copies of chromosomal regions carrying the disease allele, leading to homozygosity in this region of the genome as shown in Figure 12. Such extended regions of homozygosity are known as runs of homozygosity (ROHs). ROHs are usually between a few hunded kilobases and several mega bases long, and their length decreases with genetic distance due to recombination events (127).

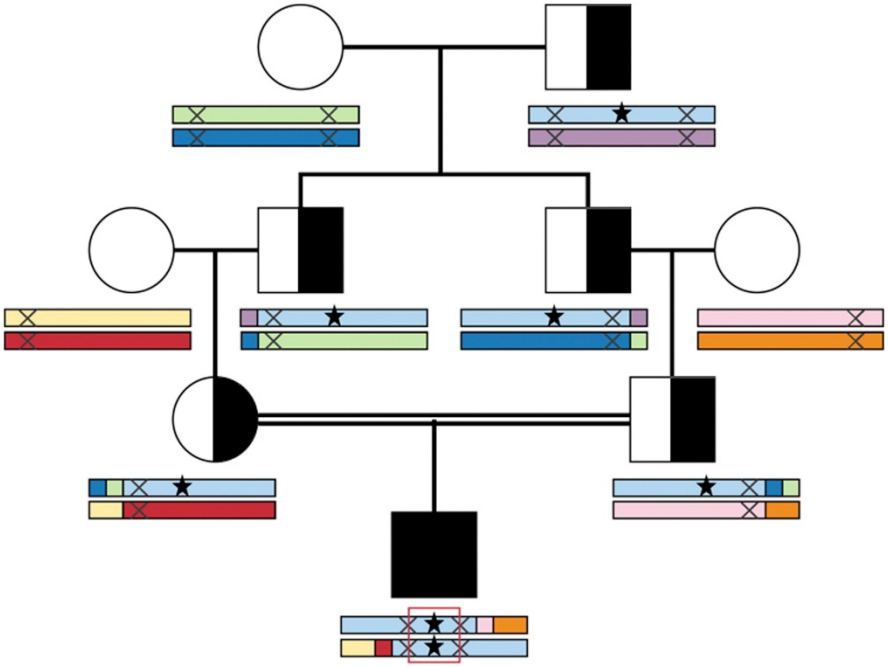


Figure 12: Segregation of a recessive disease allele in a pedigree. Reproduced from (127). The affected individual has inherited two copies of the disease allele (star). Because the allele is identical by descent, it is included in an extended region of homozygosity in the affected individual.

Genome wide association study

A popular alternative to family-based designs are GWASs. Their classical design is a case-control study, where the association of genetic markers with a trait is calculated in a large number of unrelated affected and nonaffected individuals (128). This means, an allele is associated with the trait of interest, if its frequency is significantly different between the case and the control group. Since most GWASs are based on genetic markers, an associated SNV is most likely not causative but simply in linkage disequilibrium with the true causative variant.

While the principle of a GWAS is relatively simple, there are a few statistical and design related issues that have to be taken into account (128).

First, the study has to have enough power to detect an association. In a GWAS, power is the probability of detecting a true association, or in other words the probability of a correct rejection of the null hypothesis. Several factors influence power, and not all of them can be controlled. One of the controllable factors is the sample size (129). In humans with a large genetic diversity, and especially for complex traits, a large sample size is required (130). For some monogenic traits in dogs, it has been shown that 10-20 cases and the same number of controls are sufficient to detect associations with a recessive trait, while twice as many samples were required for dominant traits (131).

In a genome-wide study, thousands of markers are analyzed simultaneously and each of these markers counts as independent test. With a significance threshold of 0.05 and a set of 500,000 markers, we would therefore expect 25,000 significant associations just by chance. To avoid such false positives, results have to be adjusted for multiple testing. A conservative multiple testing correction is the Bonferroni correction. This correction is implemented by dividing the significance threshold by the number of independent tests (markers). Bonferroni correction is however very strict and ignores that the number of genetic markers does not represent the number of independent tests due to linkage disequilibrium (128, 132). Other multiple testing correction methods include procedures based on permutation or controlling the false discovery rate (133, 134).

A confounding factor that is challenging to detect and correct for is population stratification. If the allele frequency at a given marker is different between case and control groups due to reasons other than the phenotype of interest, this results in spurious allelic associations (135). A commonly accepted way to correct for population structure, family structure and cryptic relatedness is the use of mixed models. Programs applying this method calculate a pair-wise genetic relationship matrix based on autosomal markers which is then integrated into the model as random effect (136, 137).

Next generation sequencing

After the field of DNA sequencing had advanced relatively slowly for decades, the first breakthrough was the development of the Sanger sequencing technique in 1977 (138, 139). This technique makes use of the dideoxy chain-termination method and is still used for targeted sequencing of candidate genes and validation of next generation sequencing results (140). With this first generation sequencing technique, segments of almost one kilobase can be sequenced with a high accuracy. With the development of second generation techniques

(together with the third generation also referred to as next generation sequencing) high-throughput sequencing became possible, reducing cost, time and effort (141). These techniques revolutionized genomics and enabled the use of whole genome or whole exome sequencing in research but also for diagnostic purposes (139, 142). Among the different NGS platforms that differ in factors such as speed, read length and throughput, currently the most commonly used next generation sequencing platform is Illumina (143). Illumina machines produce short read sequencing data. To date, single or paired-end reads of 150 bp in length can be obtained on Illumina instruments.

Although next generation sequencing can not detect 100 % of the genetic variation, it has generally the advantage that rare and even *de novo* variants can be detected that would be missed by GWAS with predefined marker sets. There is also a chance to at least detect some structural variants that cannot be seen in Sanger sequencing data. Furthermore, it is possible to identify causative variants in a small number of cases or even single cases. Therefore, especially in monogenic disorders caused by such rare variants, whole genome or whole exome sequencing has had a huge impact and has led to a fast increase in the number of known causative variants over the last decade (139, 144).

Aim and hypothesis of the thesis

The aim of this thesis was to identify genetic variants causing monogenic skin disorders in domestic animal species. The hypothesis was that these disorders follow a simple Mendelian inheritance.
Specifically, the genetic background of the following disorders or traits was investigated:

Cats:

1. Feline cutaneous asthenia / Ehlers-Danlos syndrome
2. Inflammatory linear verrucous epidermal nevus (ILVEN) - like lesions

Dogs:

3. Congenital cornification disorder in a Labrador Retriever and her cross-bred daughter
4. Coat colour dilution in Chow Chows
5. Ectodermal dysplasia in a Dachshund family
6. Follicular dysplasia in Curly Coated Retrievers
7. Ichthyosis in a German Shepherd
8. Lethal acrodermatitis in Bull Terriers and Miniature Bull Terriers
9. Nasal parakeratosis in Greyhounds
10. Oculocutaneous albinism in German Spitz dogs
11. Oculocutaneous albinism in a Bullmastiff

Horses:

12. Naked foal syndrome in Akhal Teke horses

To facilitate the analysis of disorders potentially related to keratins, a curated catalog of canine and equine keratin genes was created, which is presented in the last chapter of the results section.

Results

A frameshift variant in the *COL5A1* gene in a cat with Ehlers-Danlos syndrome

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	design and supervision of genetic experiments, review and editing of manuscript

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12727

A frameshift variant in the *COL5A1* gene in a cat with Ehlers-Danlos syndrome

M. Spycher*†, A. Bauer*†, V. Jagannathan*†, M. Frizzi‡, M. De Lucia‡§ and T. Leeb*†

*Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001, Bern, Switzerland. †DermFocus, University of Bern, 3001, Bern, Switzerland. ‡San Marco Veterinary Clinic,Via Sorio 114/C, 35141, Padova, Italy. §San Marco Veterinary Laboratory, Via Sorio 114/C, 35141, Padova, Italy.

Summary

Ehlers-Danlos syndrome (EDS) is a group of heritable connective tissue disorders caused by defective collagen synthesis or incorrect assembly of the collagen triple helical structure. EDS is characterised by joint hypermobility, skin hyperextensibility, abnormal scarring, poor wound healing and tissue friability. Human EDS may be caused by variants in several different genes including *COL5A1*, which encodes the collagen type V alpha 1 chain. For the present study we investigated a 1.5-year-old, spayed female, domestic shorthair cat with EDS. The affected cat showed multiple recurrent skin tears, hyperextensibility of the skin and joint abnormalities. We obtained whole genome sequencing data from the affected cat and searched for variants in candidate genes known to cause EDS. We detected a heterozygous single base-pair deletion in exon 43 of the *COL5A1* gene, namely c.3420delG. The deletion was predicted to result in a frameshift and premature stop codon: p.(Leu1141SerfsTer134). Sanger sequencing confirmed that the variant was present in the affected cat and absent from 103 unaffected cats from different breeds. The variant was also absent from a Burmese cat with EDS. Based on knowledge about the functional impact of *COL5A1* variants in other species, *COL5A1*: c.3420delG represents a compelling candidate causative variant for the observed EDS in the affected cat.

Keywords collagen, dermatology, *Felis catus*, genodermatosis, skin, whole genome sequencing

Ehlers-Danlos syndrome (EDS) is a clinically and genetically heterogeneous group of heritable connective tissue disorders. EDS is caused by defective collagen synthesis or incorrect assembly of the collagen triple helical structure. The phenotype is characterised by joint hypermobility, skin hyperextensibility, abnormal scarring, poor wound healing and tissue friability (Byers & Murray 2012; De Paepe & Malfait 2012). Collagen provides the connective tissue matrix with shape, strength and the ability to resist deformation, thus being a key structural protein of the connective tissue. Fibrillary collagen proteins consist of three either homo- or heterotrimeric polypeptide chains, designated as α chains, which together form a triple helical structure (De Paepe & Malfait 2012). As a result of defects in fibrillary collagen, the skin may become more fragile and tear more easily.

Address for correspondence

T. Leeb, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland.
E-mail: tosso.leeb@vetsuisse.unibe.ch

Accepted for publication 21 August 2018

According to the 2017 international classification there are 13 recognized subtypes of EDS in humans. Depending on the causative variant, EDS may follow an autosomal dominant or autosomal recessive mode of inheritance (Malfait *et al.* 2017).

The most common type of EDS, the so-called classical EDS (cEDS), formerly categorized as EDS I or EDS II, is caused by genetic variants in the *COL5A1* and *COL5A2* genes (Nicholls *et al.* 1996; Symoens *et al.* 2008; Malfait *et al.* 2010; Bowen *et al.* 2017). Variants in many other genes are known to have an impact on collagen structure that may also result in the clinical picture of EDS (Byers & Murray 2012; De Paepe & Malfait 2012; Malfait *et al.* 2017).

Connective tissue diseases resembling human EDS have been observed in many different mammalian species such as cattle, dogs, minks, horses, rabbits and sheep (Hegreberg *et al.* 1969; Harvey *et al.* 1990; Colige *et al.* 1999; Sequeira *et al.* 1999; Paciello *et al.* 2003; Zhou *et al.* 2012; Monthoux *et al.* 2015). A closely related phenotype is equine regional dermal asthenia (HERDA) in horses, which is caused by a variant in the *PPIB* gene (Tryon *et al.* 2007).

Furthermore, a Holstein calf with a phenotype described as a variant form of EDS and a heterozygous missense variant in the *EPYC* gene was reported (Tajima *et al.* 1999). This calf had a deficiency of dermatan sulfate proteoglycan, but the causality of the *EPYC* genetic variant was not conclusively proven. So far, neither *PPIB* nor *EPYC* variants have been reported in human EDS patients.

In cats, isolated EDS cases have been described as cutaneous asthenia or dermatosparaxis in several individual purebred and crossbred animals (OMIA 000327-9685; Scott 1974; Patterson & Minor 1977; Counts *et al.* 1980; Holbrook *et al.* 1980; Verweij & van Zuylen 1986; Plotnick *et al.* 1992; Sequeira *et al.* 1999; Benitah *et al.* 2004; Szczepanik et al. 2006; Smids 2008; Dokuzeylül *et al.* 2013; Weingart *et al.* 2014). An experimental domestic shorthair cat colony with an autosomal dominant form of EDS was established to study the biomechanical properties and wound-healing characteristics of skin (Freeman *et al.* 1989a,b). More recently, EDS has also been observed in several Burmese cats and it was suggested that an autosomal recessive form of EDS may be segregating in this breed (Hansen *et al.* 2015).

Before genetic analyses became widely available, a detailed biochemical analysis in a single affected Himalayan cat revealed a defect in ADAMTS2, the procollagen N-endopeptidase (Counts *et al.* 1980). Thus, this case corresponded to the human dermatosparaxis EDS (dEDS or formerly EDS VIIC). To the best of our knowledge, no causative genetic variant in a cat with EDS has yet been reported in the scientific literature.

In the present study, we investigated a 1.5-year-old, spayed female, domestic shorthair cat with characteristics of EDS. The affected cat showed multiple recurrent skin tears with little or no bleeding, located mainly on the dorsal neck and the shoulders, and hyperextensibility of the skin (Fig. 1). The skin extensibility index, according to Hansen *et al.* (2015), was 27%. Some of the previous lacerations had slowly healed leaving shiny alopecic scars. Other clinical findings included bilateral hip subluxation with a positive Ortolani sign even in the awake patient, bilateral carpal hyperextension with plantigrade appearance, pain and laxity during palpation of all joints and bilateral perineal hernias. The index cat was found on the street when she was a kitten together with a female littermate, which appeared to be normal at the clinical examination. Information about the parents was not available.

To confirm our hypothesis of a genetic defect related to collagen, we obtained an EDTA blood sample of the affected cat, isolated genomic DNA and performed whole genome re-sequencing at 20× coverage using 2× 150-bp reads on an Illumina HiSeq 3000 instrument. Private variants were identified by comparing the sequence from the affected cat to the feline reference genome assembly FelCat 9.0 and to genome re-sequencing data from 11 genetically diverse control cats obtained during other projects (database accessions given in Table S1). The methodology was previously described (Bauer *et al.* 2017). We identified 93 private homozygous and 2339 private heterozygous protein-changing variants (Table S2). These variants included a heterozygous frameshifting single-base deletion in exon 43 of the *COL5A1* gene [XM_023242950.1:c.3420delG, Chr13:g.93 210 344delC or XP_023098718.1:p.(Leu1141 SerfsTer134)]. None of the other private protein-changing variants were located in a known EDS candidate gene.

We confirmed the *COL5A1*:c.3420delG variant in the affected cat using Sanger sequencing and genotyped a sample of 104 genetically diverse cats including the unaffected littermate of the affected cat and a Burmese cat affected by EDS. Primers COL5A1_F1, AAGCTGGCTGAA ACCCATC and COL5A1_R1, CGAGCACTCCAGAGATGTCA were used to amplify a 418-bp amplicon containing the *COL5A1*:c.3420delG variant. Both primers were individually used as sequencing primers to obtain sequences in both orientations on an ABI 3730 capillary sequencer (Applied Biosystems). This experiment confirmed that the variant was present exclusively in the affected cat and did not occur in the other 104 genotyped cats including the Burmese cat with EDS (Fig. 2).

The identified *COL5A1*:c.3420delG variant in the affected cat causes a shift in the open reading frame, resulting in a premature stop codon, and is predicted to truncate approximately one-third of the 1837 amino acids of the wildtype *COL5A1* protein. Human and feline proteins share 96% amino acid identity. Heterozygous variants in the *COL5A1* gene may cause EDS in humans and mice (Wenstrup *et al.* 2006; Malfait *et al.* 2017). In human EDS patients, many *COL5A1* nonsense, frameshift or splice-site variants have been reported (Symoens *et al.* 2008; Malfait *et al.* 2010;

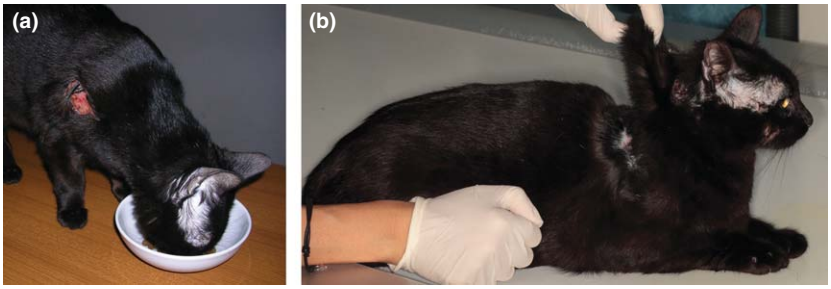


Figure 1 Ehlers-Danlos syndrome in a cat. (a) Skin tear on the right shoulder. (b) Hyperextensibility of the skin.

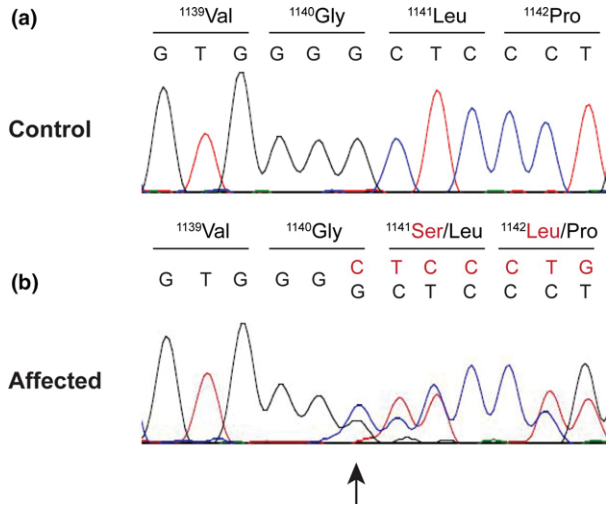


Figure 2 Electropherograms showing the *COL5A1*:c.3420delG frameshift variant. (a) Wildtype sequence of an unaffected cat. (b) Mutant sequence in the affected cat. The heterozygous deletion of a single guanine leading to a frameshift is indicated by an arrow. Wildtype sequences are shown in black, mutant sequences in red.

Byers & Murray 2012). The ClinVar database lists human pathogenic *COL5A1* frameshift variants that are comparable to the identified feline variant and lead to relatively mild clinical forms of cEDS (formerly EDS II), for example, NM_001278074.1:c.3206dup (p.Ala1070Serfs) and NM_000093.4:c.3752delC (p.Pro1251Argfs). Knowledge about the functional impact of *COL5A1* frameshift variants in humans suggests a causal role for the detected feline *COL5A1*:c.3420delG variant in the cat with EDS.

As we did not have access to RNA or protein samples from the affected cat, we could not investigate the functional consequences of the genomic variant. It seems likely that transcripts from the mutant allele are degraded by nonsense-mediated decay or other quality-control mechanisms, which might lead to a reduced amount of synthesized collagen type V alpha 1 chains and cause the clinical phenotype due to haploinsufficiency, similar to what has been observed in *Col5a1*$^{+/-}$ mice (Wenstrup *et al.* 2006). An alternative pathomechanism, in which at least some of the mutant protein is expressed and incorporated into defective collagen triple helices, also cannot be ruled out. The triple-helical structure of collagen makes it particularly sensitive to genetic variants, as each mutant protein molecule can potentially oligomerize with up to two wildtype protein molecules, thereby exerting a pronounced dominant negative effect.

No information on the parents of the affected cat was available. It is therefore impossible to investigate whether the *COL5A1*:c.3420delG variant was due to a *de novo* mutation event or whether it was actually transmitted by one of the cat's parents. Given the autosomal dominant mode of inheritance, we speculate that this variant is probably limited to the observed case.

In summary, we identified a heterozygous single-nucleotide deletion in *COL5A1* in a cat with EDS. The variant was not present in 103 unaffected control cats or a Burmese cat with EDS. The known functional impact of *COL5A1* frameshift variants in humans suggests that the detected feline *COL5A1*:c.3420delG variant is an excellent candidate causative variant for the EDS phenotype in the investigated cat.

Acknowledgements

The authors would like to thank the cat owner for donating samples and pictures. The authors also wish to thank Nathalie Besuchet Schmutz, Muriel Fragnière and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure. This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738/1).

References

Bauer A., Waluk D.P., Galichet A. *et al.* (2017) A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. *PLoS Genetics* **13**, e1006651.

Benitah N., Matousek J.L., Barnes R.F., Lichtensteiger C.A. and Campbell K.L. (2004) Diaphragmatic and perineal hernias associated with cutaneous asthenia in a cat. *Journal of the American Veterinary Medical Association* **224**, 706–9.

Bowen J.M., Sobey G.J., Burrows N.P., Colombi M., Lavallee M.E., Malfait F. & Francomano C.A. (2017) Ehlers-Danlos syndrome, classical type. *American Journal of Medical Genetics Part C (Seminars in Medical Genetics)* **175C**, 27–39.

Byers P.H. & Murray M.L. (2012) Heritable collagen disorders: the paradigm of the Ehlers-Danlos syndrome. *Journal of Investigative Dermatology* **132**, E6–11.

Colige A., Sieron A.L., Li S.W. *et al.* (1999) Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene. *American Journal of Human Genetics* **65**, 308–17.

Counts D.F., Byers P.H., Holbrook K.A. & Hegreberg G.A. (1980) Dermatosparaxis in a Himalayan cat: I. Biochemical studies of dermal collagen. *Journal of Investigative Dermatology* **74**, 96–9.

De Paepe A. & Malfait F. (2012) The Ehlers-Danlos syndrome, a disorder with many faces. *Clinical Genetics* **82**, 1–11.

Dokuzeylül B., Altun E.D., Özdogan T.H., Bozkurt H.H., Arun S.S. & Or M.E. (2013) Cutaneous asthenia (Ehlers–Danlos syndrome) in a cat. *Turkish Journal of Veterinary and Animal Sciences* **37**, 245–9.

Freeman L.J., Hegreberg G.A. & Robinette J.D. (1989a) Cutaneous wound healing in Ehlers-Danlos syndrome. *Veterinary Surgery* **18**, 88–96.

Freeman L.J., Hegreberg G.A., Robinette J.D. & Kimbrell J.T. (1989b) Biomechanical properties of skin and wounds in Ehlers-Danlos syndrome. *Veterinary Surgery* **18**, 97–102.

Hansen N., Foster S.F., Burrows A.F., Mackie J. & Malik R. (2015) Cutaneous asthenia (Ehlers-Danlos-like syndrome) of Burmese cats. *Journal of Feline Medicine and Surgery* **17**, 954–63.

Harvey R.G., Brown P.J., Young R.D. & Whitbread T.J. (1990) A connective tissue defect in two rabbits similar to the Ehlers-Danlos syndrome. *Veterinary Record* **126**, 130–2.

Hegreberg G.A., Padgett G.A., Gorham J.R. & Henson J.B. (1969) A connective tissue disease of dogs and mink resembling the Ehlers-Danlos syndrome of man. II. Mode of inheritance. *Journal of Heredity* **60**, 249–54.

Holbrook K.A., Byers P.H., Counts D.F. & Hegreberg G.A. (1980) Dermatosparaxis in a Himalayan cat: II. Ultrastructural studies of dermal collagen. *Journal of Investigative Dermatology* **74**, 100–4.

Malfait F., Wenstrup R.J. & De Paepe A. (2010) Clinical and genetic aspects of Ehlers-Danlos syndrome, classic type. *Genetics in Medicine* **12**, 597–605.

Malfait F., Francomano C., Byers P. *et al.* (2017) The 2017 international classification of the Ehlers-Danlos syndromes. *American Journal of Medical Genetics Part C (Seminars in Medical Genetics)* **175C**, 8–26.

Monthoux C., de Brot S., Jackson M., Bleul U., Walter J. (2015) Skin malformations in a neonatal foal tested homozygous positive for Warmblood Fragile Foal Syndrome. *BMC Veterinary Research* **11**, 12.

Nicholls A.C., Oliver J.E., McCarron S., Harrison J.B., Greenspan D.S. & Pope F.M. (1996) An exon skipping mutation of a type V collagen gene (*COL5A1*) in Ehlers-Danlos syndrome. *Journal of Medical Genetics* **33**, 940–6.

Paciello O., Lamagna F., Lamagna B. & Papparella S. (2003) Ehlers-Danlos-like syndrome in 2 dogs: clinical, histologic, and ultrastructural findings. *Veterinary Clinical Pathology* **32**, 13–8.

Patterson D.F. & Minor R.R. (1977) Hereditary fragility and hyperextensibility of the skin of cats: a defect in collagen fibrillogenesis. *Laboratory Investigation* **37**, 170–9.

Plotnick A., Brunt J.E. & Reitz B. (1992) Cutaneous asthenia in a cat. *Feline Practice* **20**, 9–12.

Scott D.V. (1974) Cutaneous asthenia in a cat, resembling Ehlers-Danlos syndrome in man. *Veterinary Medicine, Small Animal Clinician* **69**, 1256–8.

Sequeira L.J., Rocha N.S., Bandarra E.P., Figueiredo L.M.A. & Eugenio F.R. (1999) Collagen dysplasia (cutaneous asthenia) in a cat. *Veterinary Pathology* **36**, 603–6.

Smids M.Y. (2008) Experience with a cat with cutis asthenia (Ehlers-Danlos syndrome). *Tijdschrift Voor Diergeneeskunde* **133**, 612–4.

Symoens S., Malfait F., Renard M., André J., Hausser I., Loeys B., Coucke P. & De Paepe A. (2008) COL5A1 signal peptide mutations interfere with protein secretion and cause classic Ehlers-Danlos syndrome. *Human Mutation* **30**, E395–403.

Szczepanik M., Gołyński M., Wilkołek P., Popiel J., Śmiech A., Pomorska D. & Nowakowski H. (2006) Ehlers-Danlos syndrome (cutaneous asthenia) – a report of three cases in cats. *Bulletin of the Veterinary Institute in Pulawy* **50**, 609–12.

Tajima M., Miyake S., Takehana K., Kobayashi A., Yamato O. & Maede Y. (1999) Gene defect of dermatan sulfate proteoglycan of cattle affected with a variant form of Ehlers-Danlos syndrome. *Journal of Veterinary Internal Medicine* **13**, 202–5.

Tryon R.C., White S.D. & Bannasch D.L. (2007) Homozygosity mapping approach identifies a missense mutation in equine cyclophilin B (PPIB) associated with HERDA in the American Quarter Horse. *Genomics* **90**, 93–102.

Verweij C.G. & van Zuylen A.L. (1986) Cutaneous asthenia, a congenital skin disease in a Burmese cat. *Tijdschrift Voor Diergeneeskunde* **111**, 244–6.

Weingart C., Haußer I., Kershaw O. & Kohn B. (2014) Ehlers-Danlos like syndrome in a cat. *Schweizer Archiv für Tierheilkunde* **156**, 543–8.

Wenstrup R.J., Florer J.B., Davidson J.M., Phillips C.L., Pfeiffer B.J., Menezes D.W., Chervoneva I. & Birk D.E. (2006) Murine model of the Ehlers-Danlos syndrome. Col5a1 haploinsufficiency disrupts collagen fibril assembly at multiple stages. *Journal of Biological Chemistry* **281**, 12888–95.

Zhou H., Hickford J.G. & Fang Q. (2012) A premature stop codon in the *ADAMTS2* gene is likely to be responsible for dermatosparaxis in Dorper sheep. *Animal Genetics* **43**, 471–3.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.
Table S1 Accession numbers of 12 cat genome sequences.
Table S2 Private protein-changing variants in the sequenced cat.

Genetic variant in the *NSDHL* gene in a cat with multiple congenital lesions resembling inflammatory linear verrucous epidermal nevi (ILVEN)

Journal:	Veterinary Dermatology
Manuscript status:	published
Contributions:	design and supervision of genetic experiments, review and editing of manuscript

Veterinary Dermatology

Vet Dermatol 2018 DOI: 10.1111/vde.12699

Genetic variant in the *NSDHL* gene in a cat with multiple congenital lesions resembling inflammatory linear verrucous epidermal nevi

Michela De Lucia*†, Anina Bauer‡§, Melina Spycher‡§, Vidhya Jagannathan‡§, Erica Romano¶, Monika Welle§†† and Tosso Leeb‡§

*San Marco Veterinary Clinic, Via Sorio 114/C, 35141 Padova, Italy
†San Marco Veterinary Laboratory, Via Sorio 114/C, 35141 Padova, Italy
‡Institute of Genetics, Vetsuisse Faculty, University of Bern, Länggassstrasse 120, 3001 Bern, Switzerland
§Dermfocus, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland
¶Veterinary Hospital Gregorio VII, Piazza di Villa Carpegna 52, 00165 Rome, Italy
††Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland

Correspondence: Michela De Lucia, San Marco Veterinary Clinic, Via Sorio 114/C, 35141 Padova, Italy. E-mail: michela_delucia@virgilio.it

Background – The feline counterpart of human inflammatory linear verrucous epidermal nevus (ILVEN) has been described; however, the possible underlying developmental defect has not been investigated.

Objective – To report a case of multiple ILVEN-like lesions in a cat with a genetic variant in the *NSDHL* gene.

Animals – A 2-year-old, female, domestic short hair cat with a history of multiple alopecic, verrucous, hyperpigmented and erythematous skin lesions, following Blaschko's lines on the head, the limbs, the trunk and paw pads.

Methods and results – According to the clinical and histopathological findings, a diagnosis of multiple ILVEN-like lesions was made. Genetic investigation revealed a heterozygous missense variant in the X-chromosomal *NSDHL* gene predicted to lead to a loss-of-function of the NSDHL protein.

Conclusions and clinical importance – To the best of the authors' knowledge, this is the first case of feline ILVEN-like lesions in which a genetic cause has been proposed. Future studies to establish a causal relationship between *NSDHL* variants and skin lesions might lead to pathogenesis-directed treatments.

Introduction

Inflammatory linear verrucous epidermal nevus (ILVEN) in humans is a subgroup of epidermal nevi that in some cases can be caused by genetic variants in the *NSDHL* gene.[1] *NSDHL* variants most often lead to a syndromic condition, including congenital hemidysplasia, ichthyosiform erythroderma (nevus) and limb defects (CHILD syndrome) but in some cases only linear epidermal nevi are evident (mild CHILD syndrome).[2] Multiple skin lesions similar to ILVEN in humans have been reported in a cat although the possible genetic cause has not been investigated.[3] The aim of this case report is to describe a case of multiple verrucous cutaneous plaques, proposed to be similar to human ILVEN, in a cat with a genetic variant in the *NSDHL* gene.

Case description

A 2-year-old, female, domestic short hair cat was presented with a history of multiple skin plaques first noticed when she was approximately 1-month-old. Initially, lesions were smaller and mildly erythematous but over time they had become larger, hyperpigmented and scaly. Topical and systemic treatments, including antifungal drugs, antibiotics, glucocorticoids and ciclosporin, did not resolve the lesions. At the time of examination, multiple plaques were present on the head, one ear, chin, limbs, trunk and paw pads. They were bilaterally distributed in a non-symmetrical pattern following Blaschko's lines, and lesions almost always showed a linear configuration except on the trunk where they had a more oval shape (Figure 1). Plaques were alopecic, hyperpigmented, with an irregular verrucous surface, occasionally crusted or eroded, and encircled by a slightly erythematous margin (Figure 2). Mild to moderate pruritus was reported by the owner. Cytological examination of the skin lesions showed numerous cocci and scattered neutrophils. Dermatophytosis was ruled out based on negative fungal culture.

According to the clinical examination and the blood and urine tests, including feline immunodeficient virus (FIV) and feline leukaemia virus (FeLV) tests, the cat was otherwise healthy. Two skin biopsy specimens were taken from the axilla and the lateral thorax. These revealed a diffuse, moderate to severe epidermal hyperplasia also

Accepted 13 October 2018
Source of funding: This study was funded in part by a grant of the Swiss National Science Foundation to T.L. (CRSII3_160738).
Conflict of interest: No conflicts of interest have been declared.

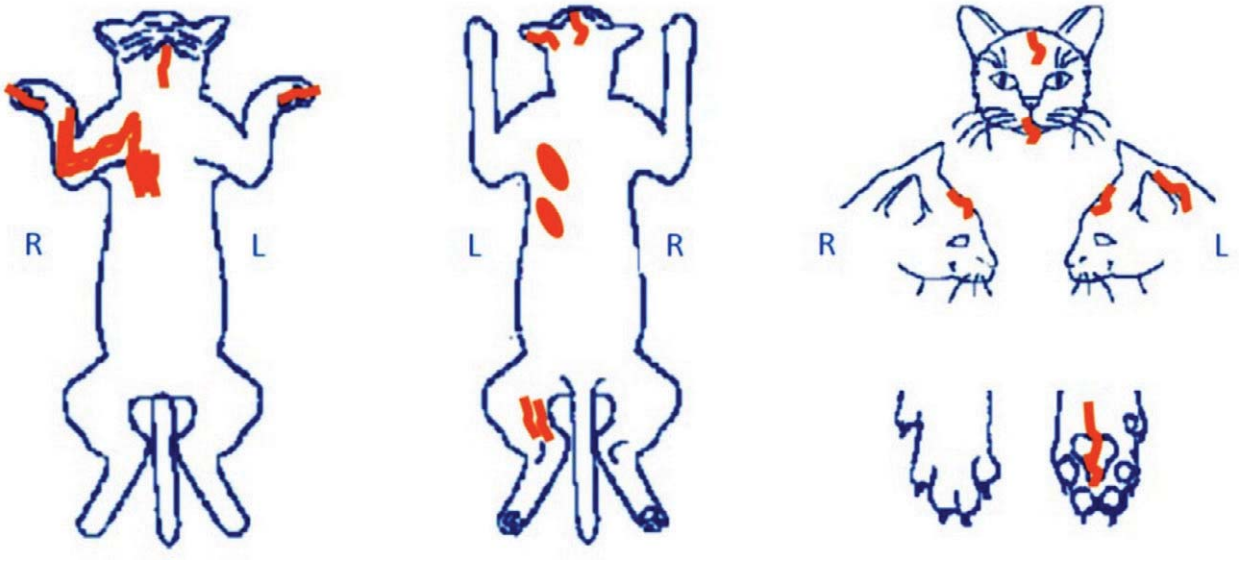


Figure 1. Illustration of the distribution of skin lesions in a cat with bilateral nonsymmetrical lesions resembling inflammatory linear verrucous epidermal nevi. L left, R right.

Figure 2. Photograph of skin lesions in a cat resembling inflammatory linear verrucous epidermal nevi.
(a) Alopecic, verrucous, hyperpigmented linear plaques following Blaschko's lines on the axilla and the medial aspect of the right forelimb. (b) A linear crusted plaque running along the midline demarcation of the chin. (c) A scaly and crusted lesion, spreading from the paw pad to the flexural aspect of the carpus in a linear fashion.

involving the infundibular walls (Figure 3). The epidermal hyperplasia was regular in the biopsy from the thorax and slightly irregular in the biopsy from the axilla. The epidermis was covered by a thick layer of mostly orthokeratotic, but also multifocal areas of parakeratotic compact keratin. Multifocally, and more pronounced in the biopsy from the axilla, there was exocytosis of inflammatory cells throughout the epidermis, small intracorneal pustules filled with neutrophils and serocellular crusts of varying size. In the surface crusts from the axilla, few cocci were visible. Sebaceous glands were moderately hyperplastic in both biopsies. In the superficial dermis, there was a moderate perivascular to interstitial infiltrate composed of mainly mast cells and fewer lymphocytes, neutrophils and melanophages. The periodic acid Schiff stain did not reveal any fungal elements. Papillomavirus-associated lesions were considered unlikely based on the absence of viral cytopathic changes. To further investigate the possible involvement of papillomavirus, PCR was performed as described previously and proved to be negative.[4]

Based on the clinical and histopathological findings, multiple plaques similar to human ILVEN were diagnosed. To probe the genetic basis of the condition, DNA from a blood sample collected in an EDTA tube was obtained. The genome of this cat was sequenced at ×31 coverage; single nucleotide variants and short indels were called with respect to the reference genome assembly FelCat 9.0 as described previously.[5] The sequencing data were deposited under project accession PRJEB7401 and sample accession SAMEA104694019 in the European Nucleotide Archive (ENA). Based on the distribution of the lesions and as the cat was female, we hypothesized

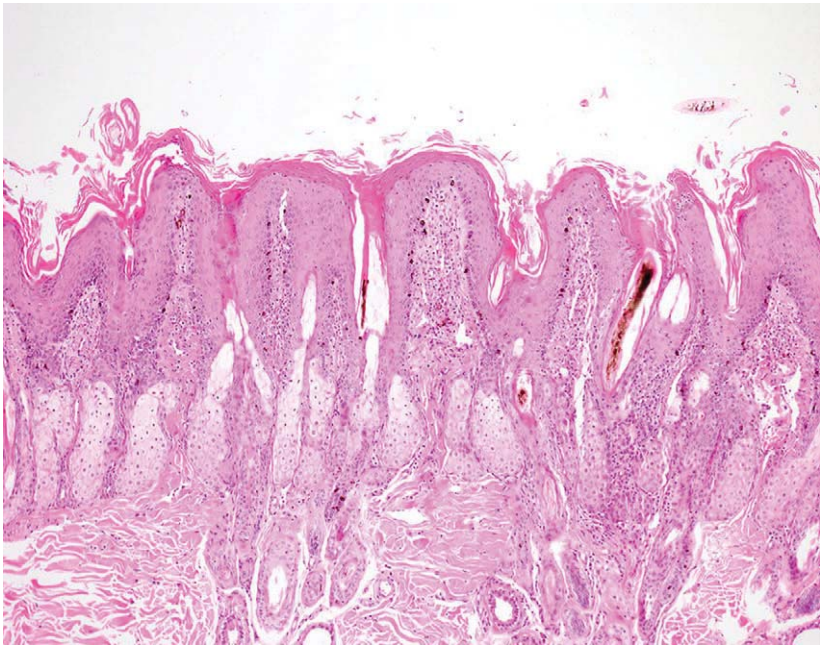


Figure 3. Photomicrograph of a skin sample from the lateral thorax of a cat with skin lesions resembling inflammatory linear verrucous epidermal nevi.
Histologically, the lesions are characterized by regular epidermal hyperplasia involving also the walls of the follicular infundibula. The epidermis is covered by a thick layer of mostly orthokeratotic, but multifocally also parakeratotic, compact keratin, which also is present within the follicular infundibula. Sebaceous glands are hyperplastic. In the superficial dermis, a moderate perivascular to interstitial infiltrate (mainly composed of mast cells and fewer lymphocytes, neutrophils and melanophages although this is not readily visible at this magnification) was present. Haematoxylin and eosin ×40.

that the condition was due to X-chromosomal functional mosaicism. We therefore filtered for variants on the X chromosome with heterozygous genotypes in the affected cat, while simultaneously having a homozygous reference (or missing) genotype in 14 additional genome sequences from unaffected control cats. This revealed a heterozygous missense variant in the *NSDHL* gene: XM_004000985.5:*c*.397A>G or XP_004001034.1:p.(Ser133Gly). The variant was absent from 93 additional genetically diverse control cats as demonstrated by Sanger sequencing. The variant was predicted to affect the essential serine residue of the catalytic tetrad 109Asn-133Ser-160Tyr-164Lys in the conserved short-chain dehydrogenase/reductase domain of the feline NSDHL protein and thus most likely leading to a functional inactivation of the protein.[6]

No treatment was prescribed and the lesions appeared unchanged after one year.

Discussion

To the best of the authors' knowledge, this is the first time that an underlying genetic cause in cats with cutaneous lesions similar to human ILVEN has been reported. According to our results, the disorder might be due to inactivating genetic variants within the *NSDHL* gene resulting in a block in cholesterol biosynthesis and the accumulation of toxic intermediates in the skin.[1] Although *NSDHL* mutations in humans frequently cause the severe syndromic condition involving limb defects (CHILD syndrome), our case did not present any relevant clinical abnormalities apart from the skin lesions and it appears to be more similar to the so-called mild CHILD syndrome, in which only skin lesions are evident. Interestingly, the only feline ILVEN case previously described presented with neurological signs including ataxia and tremors.[3] Unfortunately, in that case, genetic investigation was not performed.[3]

As distinct from the previous case,[3] our case showed a bilateral distribution of lesions. Albeit most cases of ILVEN in human are unilateral, bilateral involvement also has been reported.[1] In dogs, widespread ILVEN-like lesions have been described.[7]

Our case did not display the classical histopathological pattern of ILVEN described as alternating areas of severe parakeratotic hyperkeratosis associated with a missing stratum granulosum underneath, and areas of hyperkeratosis associated with hypergranulosis.[8] In addition, the hyperplasia of sebaceous glands has not been described before. We assume that in domestic animals, similar to what has been described in humans, variations of the histological presentation occur.[9] However, only additional cases will provide the ultimate proof that the described histological lesions are truly the consequence of a genetic variation within the *NSDHL* gene. Furthermore, we also cannot exclude the possibility that in other body locations lesions might have been more severe.

So far, the only reported treatment for ILVEN in cat is surgery.[3] However, in humans, a pathogenesis-targeted therapeutic approach to treat *NSDHL*-associated ILVEN with a combination of topical cholesterol, lovastatin and glycolic acid has shown promising results.[10]

This case report unravelled a potential inherited defect in cholesterol biosynthesis as the possible underlying pathogenesis for ILVEN-like lesions in cats. This hints at

the potential of personalized veterinary medicine and opens the door for future pathogenesis-directed treatment attempts in feline ILVEN-like cases with proven inactivating *NSDHL* genetic variants.

References

1. Bornholdt D, König A, Happle R et al. Mutational spectrum of NSDHL in CHILD syndrome. *J Med Genet* 2005; 42: e17.
2. Gantner S, Rütten A, Requena L et al. CHILD syndrome with mild skin lesions: histopathologic clues for the diagnosis. *J Cutan Pathol* 2014; 41: 787–790.
3. Sato M, Kariya K, Matsumoto M et al. Feline epidermal nevi resembling human inflammatory linear verrucous epidermal nevus. *J Vet Med Sci* 2012; 74: 1,337–1,339.
4. Munday JS, Kiupel M, French AF et al. Amplification of papillomaviral DNA sequences from a high proportion of feline cutaneous in situ and invasive squamous cell carcinomas using a nested polymerase chain reaction. *Vet Dermatol* 2008; 19: 259–263.
5. Bauer A, De Lucia M, Jagannathan V et al. A large deletion in the NSDHL gene in Labrador Retrievers with a congenital cornification disorder. *G3 (Bethesda)* 2017; 7: 3,115–3,121.
6. Filling C, Berndt KD, Benach J et al. Critical residues for structure and catalysis in short-chain dehydrogenases/reductases. *J Biol Chem* 2002; 277: 25,677–25,684.
7. White SD, Rosychuk RA, Scott KV et al. Inflammatory linear verrucous epidermal nevus in four dogs. *Vet Dermatol* 1992; 3: 107–113.
8. Patterson JWTumors of the epidermis. *Weedon's Skin Pathology*. 4th edition. Philadelphia, PA: Churchill Livingstone-Elsevier, 2016; 786.
9. Happle R, Mittag H, Küster W. The CHILD nevus: a distinct skin disorder. *Dermatology* 1995; 191: 210–216.
10. Bergqvist C, Abdallah B, Hasbani D-J et al. CHILD syndrome: a modified pathogenesis-targeted therapeutic approach. *Am J Med Genet A* 2018; 186: 733–738.

Résumé

Contexte – L'équivalent félin de l'ILVEN (Inflammatory Linear Verrucous Epidermal Nevus) de l'homme a été décrit; cependant, le développement possible de défaut sous jacent n' pas été étudié.

Objectif – Décrire un cas de lésions multiples de ILVEN-like chez un chat avec un variant génétique du gène NSDHL.

Sujet – Un chat européen femelle de 2 ans, avec des antécédents de lésions cutanées érythémateuses, hyperpigmentées, verruqueuses et alopéciques multiples, localisées sur les lignes de Blaschko sur la tête, les membres, le tronc et les coussinets.

Méthodes et résultats – Selon les données cliniques et histopathologiques, un diagnostic de lésions ILVEN-like multiple a été posé. Les recherches génétiques ont révélé un variant contre-sens hétérozygote du gène NSDHL du chromosome X prédisposant à une perte de fonction de la protéine BSDHL.

Conclusions et importance clinique – À la connaissance des auteurs, ceci est le premier cas de lésions ILVEN-like félin avec une cause génétique proposée. D'autres études sont nécessaires pour établir la relation entre les variants de NSDHL et les lésions cutanées pouvant mener à des traitements pathogéniques.

Resumen

Introducción – se ha descrito el equivalente felino del nevus epidérmico verrucoso lineal inflamatorio humano (ILVEN); sin embargo, el posible defecto subyacente en su desarrollo no ha sido investigado.

Objetivo – reportar un caso de lesiones similares a ILVEN en un gato con una variante genética en el gen *NSDHL*.

Animales – un gato doméstico de 2 años de edad, con pelo corto y con antecedentes de múltiples lesiones cutáneas alopécicas, verrugosas, hiperpigmentadas y eritematosas, siguiendo las líneas de Blaschko en la cabeza, las extremidades, el tronco y las almohadillas de las patas.

Métodos y resultados – de acuerdo con los hallazgos clínicos e histopatológicos, se realizó un diagnóstico de múltiples lesiones similares a ILVEN. La investigación genética reveló una variante heterocigótica heterocigótica en el gen *NSDHL* del cromosoma X que se asocia a una pérdida de función de la proteína NSDHL.

Conclusiones e importancia clínica – a entender de los autores, este es el primer caso de lesiones felinas similares a ILVEN en las que se ha propuesto una causa genética. Los estudios futuros para establecer una relación causal entre las variantes de *NSDHL* y las lesiones cutáneas podrían conducir a tratamientos dirigidos por la patogénesis.

Zusammenfassung

Hintergrund – Das feline Gegenstück zum entzündlichen linearen verrukösen epidermalen Nevus (ILVEN) des Menschen ist beschrieben worden; es wurde jedoch der mögliche zugrundeliegende Entwicklungsdefekt noch nicht untersucht.

Ziel – Der Bericht eines Falles mit multiplen ILVEN-ähnlichen Läsionen bei einer Katze mit einer genetischen Variante des NSDHL Gens.

Tiere – Eine 2 Jahre alte weibliche europäische Kurzhaarkatze mit einer Anamnese von multiplen haarlosen, verrukösen, hyperpigmentierten und erythematösen Hautveränderungen, die den Blaschko Linien am Kopf, an den Extremitäten, dem Rumpf und den Pfotenballen folgten.

Methoden und Ergebnisse – Den klinischen und histopathologischen Ergebnissen zufolge wurde die Diagnose einer multiplen ILVEN-ähnlichen Läsion gestellt. Die genetische Untersuchung ergab eine

Missense (Fehlsinn) Variante im X-chromosomalen NDSHL Gen, wodurch vorhergesagt werden konnte, dass das NSDHL-Protein seine Funktion verlieren würde.
Schlussfolgerungen und klinische Bedeutung – Nach bestem Wissen der Autoren handelt es sich hierbei um den ersten Fall einer ILVEN-ähnlichen Veränderung bei der Katze, bei dem eine genetische Ursache angenommen wird. Zukünftige Studien sollten einen kausalen Zusammenhang zwischen NSDHL Varianten und Hautveränderungen herstellen, was zu einer der Pathogenese entsprechenden Behandlung führen könnte.

要約
背景 – 人の炎症性線状疣贅状表皮母斑(ILVEN)に相当する疾患が猫において報告されている。しかし、根底にある可能性のある発達上の欠陥は調査されていない。
目的 – 本研究の目的は、NSDHL遺伝子に遺伝子変異を認め、複数のILVEN様病変を有する猫の1例を報告することである。
被験動物 – 頭部、四肢、体幹および肉球のBlaschko線に沿って、複数の脱毛症、疣贅、色素沈着および紅斑性病変の病歴を有する2歳の雌の雑種猫。
方法および結果 – 臨床および組織病理学的所見にしたがって、複数のILVEN様病変と診断した。遺伝子解析により、NSDHLタンパクに機能欠損をもたらすと予測されるX染色体NSDHL遺伝子のヘテロミスセンス変異体が明らかになった。
結論と臨床的重要性 – 著者の知る限りにおいて、本症例は遺伝的な原因が提唱されたネコILVEN様病変の最初の症例である。 NSDHL変異体と皮膚病変との間の因果関係を確立するための今後の研究は、病原性に対する治療を導く可能性がある。

摘要
背景 — 已报道猫具有与人类炎性线性疣状表皮痣(ILVEN)相似的疾病；然而,可能的潜在发育缺陷尚未被研究。
目的 — 报告一例NSDHL基因遗传性变异的猫病例,皮肤具有多处ILVEN样病变。
动物 — 一只2岁、雌性、短毛家猫,有多发性脱发、疣状、色素沉着和红斑性皮肤病变的病史,沿Blaschko's线分布在头部、四肢、躯干和爪垫上。
方法和结果 — 根据临床和组织病理学结果,对多个ILVEN样病变进行诊断。遗传学研究揭示,在X染色体NSDHL基因中,存在一个杂合错义变体,推测会导致NSDHL蛋白功能丧失。
结论和临床价值 — 据作者所知,这是第一例猫发生ILVEN样病变,其中遗传原因已被提出。需要更多研究去建立NSDHL变体与皮肤病变之间的因果关系,以实现对因治疗。

Resumo
Contexto – A versão felina do nevo epidérmico verrucoso inflamatório linear humano (NEVIL) já foi descrita; entretanto, o possível defeito de desenvolvimento de base ainda não foi investigado.
Objetivo – Relatar um caso de lesões NEVIL-símile múltiplas em uma gata com uma variação genética no gene *NSDHL*.
Animais – Uma gata doméstica de pelo curto de dois anos de idade, com um histórico de múltiplas lesões cutâneas alopécicas, hiperpigmentadas e eritematosas, seguindo as linhas de Blaschko na cabeça, membros, tronco e coxins.
Métodos e resultados – O diagnóstico de lesões cutâneas NEVIL-símile foi realizado baseado nos achados clínicos e histopatológicos. A pesquisa genética revelou mutação não-sinônima heterozigótica no gene *NSDHL* do cromossomo X geradora de perda de função da proteína NSDHL.
Conclusões e importância clínica – De acordo com os conhecimentos dos autores, este é o primeiro caso de lesões NEVIL-símile em felinos em que uma causa genética tenha sido proposta. Estudos futuros para estabelecer uma relação causal entre as variantes de *NSDHL* e lesões cutâneas podem direcionar tratamentos específicos para a patogenia da doença.

A large deletion in the *NSDHL* gene in Labrador Retrievers with a congenital cornification disorder

Journal:	G3: Genes, Genomes, Genetics
Manuscript status:	published
Contributions:	genetic analyses, illustrations, original draft, review and editing of manuscript

G3 Genes | Genomes | Genetics

INVESTIGATION

A Large Deletion in the *NSDHL* Gene in Labrador Retrievers with a Congenital Cornification Disorder

Anina Bauer,*,† Michela De Lucia,‡ Vidhya Jagannathan,*,† Giorgia Mezzalira,‡ Margret L. Casal,§ Monika M. Welle,†, and Tosso Leeb*,†,1**

*Institute of Genetics, †Dermfocus, and **Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3001, Switzerland, ‡San Marco Veterinary Clinic and Laboratory, Padova 35141, Italy, §Section of Medical Genetics, Department of Clinical Studies, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104

ORCID IDs: 0000-0002-6765-4012 (A.B.); 0000-0003-0553-4880 (T.L.)

ABSTRACT In heterozygous females affected by an X-linked skin disorder, lesions often appear in a characteristic pattern, the so-called Blaschko's lines. We investigated a female Labrador Retriever and her crossbred daughter, which both showed similar clinical lesions that followed Blaschko's lines. The two male littermates of the affected daughter had died at birth, suggesting a monogenic X-chromosomal semidominant mode of inheritance. Whole genome sequencing of the affected daughter, and subsequent automated variant filtering with respect to 188 nonaffected control dogs of different breeds, revealed 332 heterozygous variants on the X-chromosome private to the affected dog. None of these variants was protein-changing. By visual inspection of candidate genes located on the X-chromosome, we identified a large deletion in the *NSDHL* gene, encoding NAD(P) dependent steroid dehydrogenase-like, a 3β-hydroxysteroid dehydrogenase involved in cholesterol biosynthesis. The deletion spanned >14 kb, and included the last three exons of the *NSDHL* gene. By PCR and fragment length analysis, we confirmed the presence of the variant in both affected dogs, and its absence in 50 control Labrador Retrievers. Variants in the *NSDHL* gene cause CHILD syndrome in humans, and the bare patches (*Bpa*) and striated (*Str*) phenotypes in mice. Taken together, our genetic data and the known role of *NSDHL* in X-linked skin disorders strongly suggest that the identified structural variant in the *NSDHL* gene is causative for the phenotype in the two affected dogs.

KEYWORDS
dog
canis lupus familiaris
X-chromosome
dermatology
skin
lines of Blaschko
whole genome sequencing
animal model

A typical clinical sign of X-linked dominant or semidominant genodermatoses is the characteristic skin patterning in heterozygous females. The mechanism behind these patterns is the random X-chromosome inactivation (also called lyonization) during early embryonic development. In heterozygous females, cells with the inactivated X-chromosome carrying the pathogenic variant give rise to normal skin, whereas inactivation of the wild-type X-chromosome results in skin lesions. This leads to a visible functional mosaicism with patches of normal or lesioned skin following the lines of Blaschko (Happle 2006; Happle 2016). In humans, examples for functional X-chromosome mosaicism include disorders such as X-linked dominant chondrodysplasia punctate caused by variants in the *EBP* gene (Braverman *et al.* 1999; Derry *et al.* 1999), focal dermal hypoplasia with causal variants in the *PORCN* gene (Grzeschik *et al.* 2007; Wang *et al.* 2007), incontinentia pigmenti caused by impaired NF-kB activation due to variants in *IKBKG* (Smahi *et al.* 2000), and the IFAP syndrome known to be caused by variants in *MBTPS2* encoding a zinc metalloprotease (Oeffner *et al.* 2009). Many of these disorders are lethal in hemizygous males. However, disorders that are sublethal or nonlethal in male patients, such as Menkes disease or X-linked dyskeratosis congenita caused by variants in the *ATP7A* and *DKC1* genes, respectively, also exist (Kaler *et al.* 1994; Heiss *et al.* 1998; Happle 2016).

Skin lesions following Blaschko's lines in heterozygous females are also known in animals with X-linked heritable phenotypes, for example, incontinentia pigmenti and brindle 1 in horses (Towers *et al.* 2013; Murgiano *et al.* 2016), streaked hairlessness in cattle (Murgiano *et al.*

doi: https://doi.org/10.1534/g3.117.1124
Manuscript received June 22, 2017; accepted for publication July 15, 2017; published Early Online July 21, 2017.

Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.1124/-/DC1.
[1]Corresponding author: Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109A, 3001 Bern, Switzerland. E-mail: tosso.leeb@vetsuisse.unibe.ch

2015), or X-linked hypohidrotic ectodermal dysplasia in dogs (Casal *et al.* 1997).

A unique patterning of the skin occurs in girls with CHILD syndrome (congenital hemidysplasia with ichthyosiform nevus and limb defects, OMIM #308050; Happle *et al.* 1980). This rare X-linked semidominant disorder is caused by heterozygous genetic variants in the *NSDHL* gene, encoding NAD(P) dependent steroid dehydrogenase-like, also termed sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating (EC:1.1.1.170), a 3β-hydroxysteroid dehydrogenase involved in cholesterol biosynthesis (König *et al.* 2000). These variants are lethal in hemizygous males (Happle *et al.* 1980). The mutational spectrum of *NSDHL* in CHILD syndrome is broad, and includes missense variants exchanging conserved amino acids in the encoded protein, as well as deletions, insertions or splice site variants and the complete deletion of the gene. As different variants cause a clinically comparable phenotype, it is suggested that they lead to a loss of function in a critical step within the cholesterol biosynthesis pathway, block the synthesis of cholesterol and lead to an aggregation of toxic intermediates (Bornholdt *et al.* 2005; Kim *et al.* 2005). The exact mechanism causing the CHILD phenotype is, however, not known (Kim *et al.* 2005). Clinically, CHILD syndrome is heterogeneous with involvement of different organs, variation in severity, and different degrees of limb hypoplasia. So far, no clear correlation was found between the nature and location of the genetic variants in the *NSDHL* gene and the severity or the differences in clinical signs (Bornholdt *et al.* 2005).

In the present study, we investigated a female Labrador Retriever and her crossbred daughter, who both showed characteristic skin lesions following Blaschko's lines. Earlier reports described similar, but not identical histologic lesions in six Rottweiler dogs, one Siberian Husky, and five female Labrador Retrievers of American and Canadian origin (Lewis *et al.* 1998; Scott and Miller 2000; Hargis *et al.* 2013). All dogs were female, but the molecular etiology was not investigated in any of these cases. The aim of the present study was to identify the causative genetic variant for the phenotype in the affected Labrador Retriever of European origin and her crossbred daughter.

MATERIALS AND METHODS

Ethics statement

All animal experiments were performed according to the local regulations. The dogs in this study were examined with the consent of their owners. The study was approved by the "Cantonal Committee for Animal Experiments" (Canton of Bern; permits 75/16 and 38/17).

Clinical examination

A 7-month-old female crossbred dog was presented for a pruritic generalized scaling dermatitis, offensive odor, and severe lameness. Skin lesions had been observed soon after birth. General physical examination and a thorough dermatological workup, including skin cytology, microscopic examination of the hairs, and multiple deep skin scrapings were performed. Complete blood cell count, serum biochemistry panel, urinalysis, hemostatic profile, thyroid hormones, and thyroid stimulating hormone (TSH) measurement were also conducted. Multiple skin biopsies were obtained for histopathologic evaluation.

According to the owners and the referring veterinarian, the dog's mother, a purebred Labrador Retriever, had been affected by the same dermatological lesions since she was a puppy, and had been biopsied a few years previously. Moreover, the two male siblings of the female index patient died soon after birth. As those data raised the suspicion of a congenital hereditary disease, ethylenediaminetetraacetic acid (EDTA) anti-coagulated blood samples from both the mother and the daughter were collected for genetic investigations.

Skin biopsies and histopathological examination

A total of four 6-mm punch biopsies were taken from affected skin of the daughter. Three biopsies were taken at the initial visit from dorsal neck, right lateral elbow, and right tibial region. One biopsy was taken 1 yr later from the dorsal neck. All biopsies were taken under local anesthesia and immediately fixed in 10% buffered formalin, embedded in paraffin, cut as 4 μm sections, and stained with hematoxylin and eosin prior to the histological evaluation. Skin formalin-fixed and paraffin-embedded samples from the mother were retrieved from the laboratory archive for histopathological revision.

DNA isolation

Genomic DNA was isolated from EDTA blood of the two affected dogs using a Maxwell RSC instrument (Promega). We additionally used DNA from EDTA blood of 29 female and 21 male nonaffected control Labrador Retrievers of European origin that had been stored in the Vetsuisse Biobank, and two female American Labrador Retrievers affected by follicular parakeratosis, which were described earlier (Hargis *et al.* 2013).

Whole genome resequencing, SNP, and short indel calling

For resequencing of one of the affected dogs, we prepared a PCR-free genomic fragment library with 350 bp insert size, and collected roughly 36× coverage data on an Illumina HiSeq3000 instrument (2 × 150 bp). Read mapping, aligning and variant calling was done as described before (Bauer *et al.* 2017). Briefly, sequence reads were mapped to the dog reference genome CanFam 3.1, and aligned using Burrows-Wheeler Aligner (BWA) version 0.7.5a (Li and Durbin 2009) with default settings. The output SAM file was converted to BAM, and the reads sorted by chromosome using Samtools (Li 2011). PCR duplicates were marked using Picard tools (http://sourceforge.net/projects/picard/). To perform local realignments, and to produce a cleaned BAM file, the Genome Analysis Tool Kit (GATK version 2.4.9, 50) was used. Putative SNVs were identified using GATK HaplotypeCaller in gVCF mode, and subsequently genotyped per-chromosome and genotyped across all samples simultaneously (McKenna *et al.* 2010). Filtering was performed using the variant filtration module of GATK. To predict the functional effects of the called variants, SnpEFF (Cingolani *et al.* 2012) software, together with NCBI annotation release 103 for CanFam 3.1, was used. For variant filtering we used 188 control genomes, which were either publicly available (Bai *et al.* 2015) or produced during other projects of our group. A list of these control genomes is given in Supplemental Material, Table S1.

Structural variant detection

Functional candidate genes located on the X chromosome were visually inspected for structural variants using the BAM file from the affected dog and the Integrative Genomics Viewer (Robinson *et al.* 2011). The selection of candidate genes was based on known X-linked human genodermatoses (Lemke *et al.* 2014), and included *ATP7A*, *DKC1*, *EBP*, *EDA*, *EFNB1*, *IKBKG*, *MBTPS2*, *NSDHL*, *PORCN*, *SAT1*, and *STS*.

PCR, fragment length analysis, and Sanger sequencing

To confirm the presence of the large heterozygous deletion in the two affected dogs, and its absence in the nonaffected control dogs, we performed a long-range PCR using the three primers NSDHL_F: TGCCATGAACATCTGGAGAG, NSDHL_R1:

Figure 1 Clinical phenotype. (A) Affected daughter at 7 months of age. (B) Lesions following Blaschko's lines. (C) Hyperkeratosis of paw pads. (D) Cluster of dilated follicular ostia.

ACCCCAAACAACGAATCCT, NSDHL_R2: ACAGCTTCCCCTGC TAAGGT, and SequalPrep long range polymerase (Thermo Fisher). In heterozygous dogs, this resulted in PCR products with sizes of 753 bp for the wildtype and 1166 bp for the deletion allele. The product of the primers NSDHL_F and NSDHL_R1 flanking the deletion on the wild-type allele was too long to be amplified using these PCR primers (15,565 bp). The amplified products were analyzed using a Fragment Analyzer capillary electrophoresis instrument (AATI).

We directly sequenced the PCR products to confirm their identity on an ABI 3730 capillary sequencer (Life Technologies) after treatment with exonuclease I and shrimp alkaline phosphatase. The sequence data were analyzed using Sequencher 5.1 (GeneCodes).

Gene analysis

We used the CanFam 3.1 reference genome assembly for all analyses. The numbering within the canine *NSDHL* gene corresponds to the transcript with the accession XM_014111859.1, and its predicted translated protein with the accession XP_013967334.1.

Data availability

An IGV screenshot illustrating the large genomic deletion is shown in Figure S1. Whole genome sequencing data of the affected crossbred dog and control dogs used for private variant filtering are listed in Table S1. Control Labrador Retrievers are listed in Table S2. Whole genome sequence data of the affected dog were deposited at the European Nucleotide Archive (ENA, project accession PRJEB16012, sample accession SAMEA104125075).

RESULTS

Qualitative phenotype description

No abnormalities except a stunted growth were found on the affected daughter during general physical examination. Linear hyperplastic and partially alopecic lesions, covered with thick brown scales and clusters of dilated follicular ostia, were the most prominent dermatological features (Figure 1). The lesions were distributed along Blaschko's lines in a bilateral rather symmetrical fashion, and were more evident on the limbs, the head, the neck, and the dorsal trunk. The abdominal and inguinal skin appeared normal. Frond-like hyperkeratotic lesions at the margin of all the pawpads with occasional horn-like projections were considered the most probable cause of the visible lameness. Cytological examination of the linear hyperplastic lesions revealed the presence of variable numbers of coccoid bacteria, and a large number of *Malassezia* yeasts, which

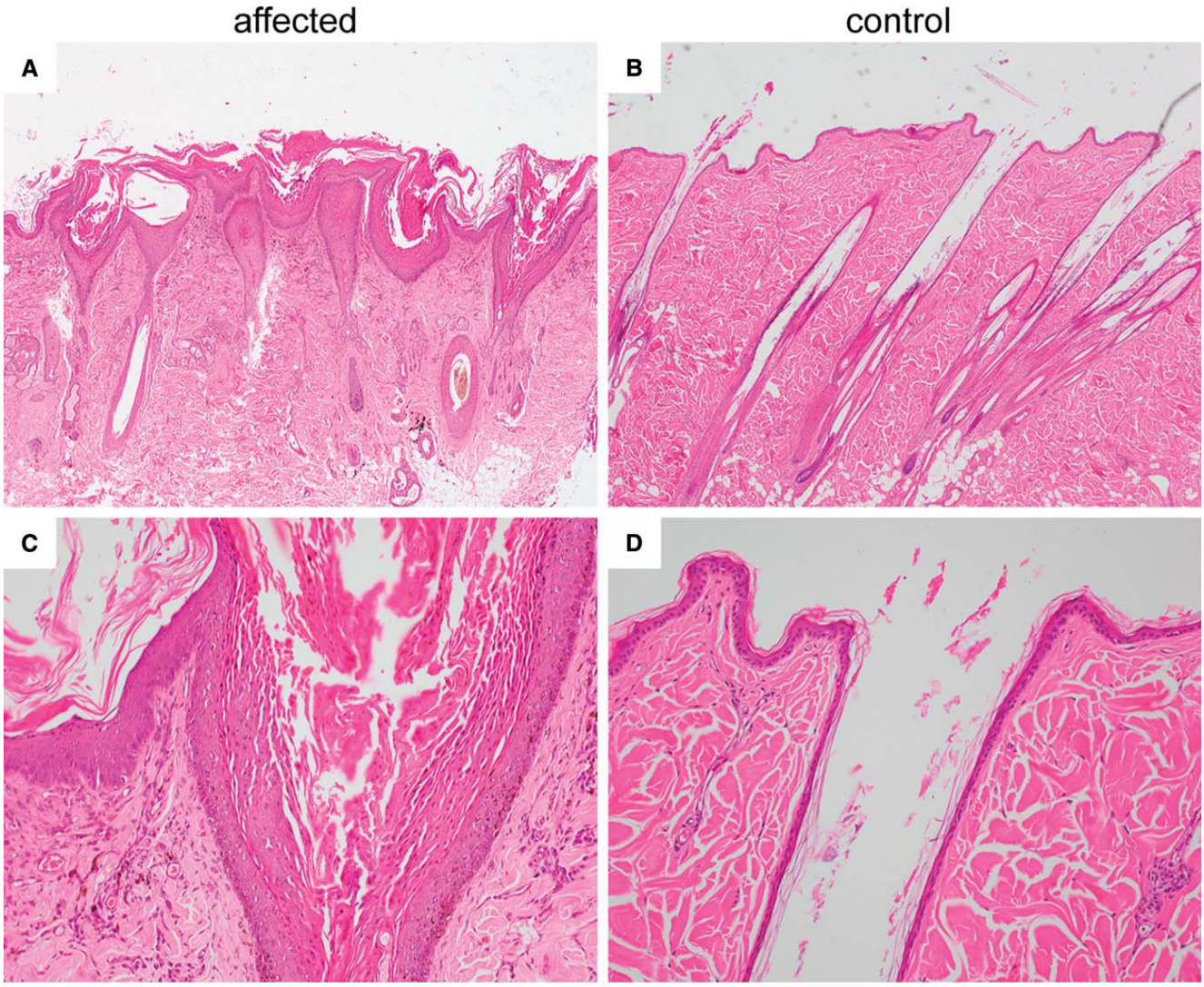


Figure 2 Histopathologic findings in an affected *vs.* control dog. (A) Photomicrograph of a skin biopsy of the affected daughter at 7 months of age depicting a moderately hyperplastic epidermis and severely hyperplastic infundibular epithelium. The infundibuli are filled with densely packed parakeratotic keratin, which is protruding above the epidermal surface. The interfollicular epidermis is covered by moderate to large amounts of laminar to compact, sometimes orthokeratotic, sometimes parakeratotic, keratin. Within the dermis pigmentary, incontinence and a moderate perivascular infiltrate is present. Hematoxylin and Eosin 40×. (B) Skin of a nonaffected dog with normal thickness of the epidermis and infundibular walls. The epidermis is covered by basket-weave orthokeratotic keratin, and the infundibuli are filled with a small amount of orthokeratotic infundibular keratin. Hematoxylin and Eosin 40×. (C, D) Skin sections of the same dogs as in (A) and (B) at higher magnification. Note the severe parakeratotic hyperkeratosis in the infundibulum of the affected dog whereas the neighboring epidermis is covered by orthokeratotic keratin. Hematoxylin and Eosin 200×.

were suspected to substantially contribute to the pruritus and the offensive odor. Results of the blood tests and urinalysis were unremarkable.

Histopathological examination

The histopathological findings were identical in all biopsies (Figure 2). Multifocally, the epidermis and the wall of the hair follicular infundibuli were moderately to severely hyperplastic, with abrupt transition to normal skin. Within the hyperplastic area, the infundibular epithelium was covered by thick layers of densely packed parakeratotic keratin, which was distending the infundibuli. The parakeratotic keratin was often protruding above the epidermal surface. The size of the keratohyalin granules within the granular cell layers of the epidermis and the infundibular wall was within the normal range. Within the parakeratotic keratin, multifocally variable numbers of coccoid bacteria were present, and occasionally the lumen of infundibuli contained degenerate neutrophils. Sebaceous glands appeared normal. The interfollicular epidermis was covered by moderate to large amounts of laminar to compact mostly orthokeratotic, but also some parakeratotic keratin. Within the keratin layers of the epidermis, multifocally degenerate neutrophils, nuclear debris, and small numbers of coccoid bacteria were present. Multifocally exocytosis of neutrophils was seen. Within the superficial dermis there was a mild pigmentary incontinence and a moderate perivascular infiltrate composed of neutrophils, mast cells, and fewer lymphocytes.

Table 1 Single nucleotide and small indel variants detected by whole genome resequencing

Filtering Step	Variants[a]
Heterozygous variants in whole genome	979,328
Heterozygous variants on X chromosome	25,986
Private heterozygous variants on X chromosome	332
Protein-changing private variants on X chromosome	0

[a] Only variants that passed the GATK quality thresholds are reported.

Whole genome resequencing

Given the characteristic and comparable skin lesions in the affected dogs, mother, and daughter, which followed Blaschko's lines, and the perinatal death of the male littermates, we hypothesized that the mode of inheritance was monogenic X-linked semidominant. We resequenced the genome of the affected daughter at 36× coverage, and called SNVs and short indels with respect to the canine reference genome assembly CanFam 3.1. We then searched for heterozygous variants in the genome sequence of the affected dog that were not present in 188 control dogs of different breeds. We found 332 heterozygous variants on the X-chromosome that were absent from the control dogs. However, none of these variants was predicted to be protein-changing variant (Table 1).

Identification of the causative variant

Given that the automated pipeline did not detect any private protein changing variants on the X-chromosome, we hypothesized that a larger structural variant might be causative for the disorder. Structural variants such as large insertions, deletions, duplications, or inversions would be missed by the applied variant detection software. We therefore selected all 11 known X-chromosomal genes involved in human genodermatoses as functional candidate genes, and visually inspected them for structural variants (Lemke *et al.* 2014). These functional candidate genes were: *ATP7A*, *DKC1*, *EBP*, *EDA*, *EFNB1*, *IKBKG*, *MBTPS2*, *NSDHL*, *PORCN*, *SAT1*, and *STS*.

In the chromosomal region of the *NSDHL* gene, a large structural variant was detected in heterozygous state in the genome of the affected dog. It was a deletion spanning 14,399 bp, including the last three exons of the *NSDHL* gene (Figure S1). The formal variant designation is chrX:120,749,179_120,763,577del14,399. The deletion truncates 192 (53%) of the 361 codons of the wildtype canine reading frame. No structural variants affecting the coding regions of the other candidate genes were detected.

Fragment length analysis and Sanger sequencing

To confirm the presence of the deletion in both affected dogs, a PCR approach was chosen. We designed two primers flanking the large deletion as well as one primer inside the deletion, and performed a long-range PCR with these three primers. In the two affected dogs, two PCR products resulted, which were consistent with the expected fragment sizes in a heterozygous genotype (Figure 3). In 50 control Labrador Retrievers, only the smaller wildtype band was detected, as expected for dogs not carrying the deletion. In addition, a PCR with only the primers outside the deletion was performed, and the resulting PCR products in both cases were Sanger sequenced. The electropherograms of the variant allele confirmed the 14,399 bp deletion.

DISCUSSION

In the present study, we identified a large deletion in the *NSDHL* gene in two female dogs whose pedigree and clinical signs suggested an X-linked semidominant disorder. The ~14 kb deletion included the last three exons of the *NSDHL* gene, and was present in the heterozygous state in the two affected dogs, but absent in 50 nonaffected Labrador Retrievers. The deletion truncated more than half of the open reading frame including the codons for the single transmembrane domain that anchors NSDHL in the ER membrane (Caldas and Herman 2003). Therefore, it seems unlikely that any protein that might be putatively expressed from the mutant allele would be functional.

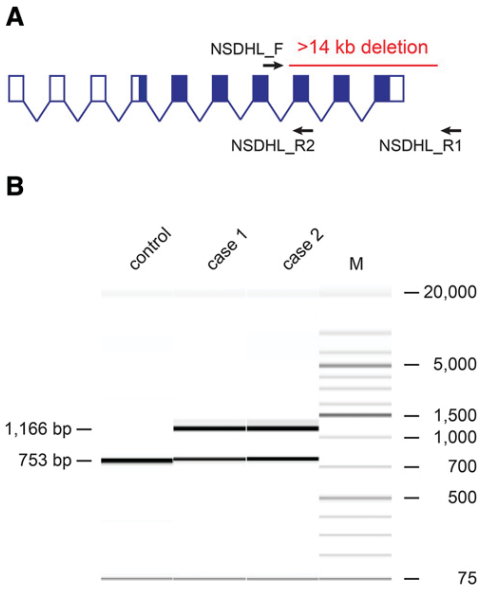


Figure 3 Confirmation of the deletion by PCR. (A) A PCR with the three primers NSDHL_F1, NSDHL_R1 and NSDHL_R2 was performed to genotype cases and controls by fragment length analysis. The exon and intron sizes of the canine *NSDHL* gene are not drawn to scale. The protein-coding region is indicated by solid filling of the exons. Note that the number of 5′-untranslated exons varies between species and transcript isoforms, whereas the seven protein-coding 3′-exons are highly conserved between human, mouse, and dog. (B) A control dog homozygous for the wildtype allele showed only a single band of 753 bp generated by amplification with the primers NSDHL_F and NSDHL_R2. The two cases heterozygous for the deletion showed the wildtype band, and an additional 1166 bp band that resulted from the amplification of NSDHL_F and NSDHL_R1 on the deletion allele. The primers NSDHL_F and NSDHL_R1 did not amplify any specific product on the wildtype allele as their binding sites were >15 kb apart.

In humans, the mutational spectrum of *NSDHL* is broad (Bornholdt *et al.* 2005). One of the known human *NSDHL* variants involves a deletion of the three last coding exons, similar to the canine deletion reported in our study. The girl carrying this deletion had CHILD syndrome, with an inflammatory epidermal nevus affecting the left side of the body. She had oligodactyly, with only three fingers on the left hand and one toe on the left foot (Kim *et al.* 2005). As the name of the human disorder suggests, such limb defects were seen in most patients with CHILD syndrome, albeit with varying severity and location (Bornholdt *et al.* 2005).

The phenotype of the dogs was related, but not identical to the clinical signs of CHILD syndrome in humans. The skin lesions in the two affected dogs followed the lines of Blaschko, but, interestingly, they did not show the typical strict lateralization of the inflammatory nevus seen in almost all human cases (Bornholdt *et al.* 2005). However, in human CHILD syndrome, bilateral involvement and very mild phenotypes without limb deformities have also been reported (König *et al.* 2002; Bittar *et al.* 2006).

Variants in the murine *Nsdhl* gene cause the bare patches (*Bpa*) and striated (*Str*) phenotypes in mice. Characteristic for heterozygous *Bpa* females is the hyperkeratotic skin eruption 5–7 d after birth. The lesions leave bare patches arranged in a typical striped pattern after resolving. Mice with this phenotype may also be affected with skeletal dysplasia, cataracts, and microphthalmia, and they are typically smaller than nonaffected littermates. *Str* heterozygous females show a milder phenotype, which is characterized by a striped coat, visible 12–14 d after birth (Liu *et al.* 1999). In both mouse mutants, limb reduction defects and diffuse skin lesions restricted to one side of the body were not seen, which is similar to the dogs described in the present study, but differing from the human phenotype (Caldas *et al.* 2005). In these mouse models, it was also found that the labyrinth layer of the fetal placenta in affected male embryos was always thinner, and fewer fetal vessels were present, possibly leading to death in midgestation (Caldas *et al.* 2005).

Interestingly, the histopathology of the affected dog skin from our study was very similar to the histopathologic findings in six female American Rottweilers and one female Siberian Husky published earlier (Lewis *et al.* 1998; Scott and Miller 2000). Different from the histology described in these dogs, the interfollicular epidermis in our cases was covered with mostly orthokeratotic keratin, and the parakeratosis was mainly restricted to the infundibulum. Clinically, the main differences between the Labrador Retrievers in the present study and the cases described earlier were the severe involvement of the pawpads and the lack of other noncutaneous congenital abnormalities apart from stunted growth.

In addition, follicular parakeratosis has been reported in five Labrador Retrievers (Hargis *et al.* 2013). These five dogs were also all female and showed multifocal crusted papules and plaques, follicular casts, comedones, and hair loss. In contrast to the phenotype in the two cases from our study, no involvement of nonhaired skin was reported. Interestingly, mural folliculitis and apoptotic cells in the hair follicle infundibili, two prominent histological features in the five dogs described by Hargis *et al.* (2013), were not observed in the two cases described here.

We were able to obtain biobanked DNA samples from two of the five Labrador Retrievers with follicular parakeratosis described by Hargis *et al.* (2013). These dogs did not carry the *NSDHL* deletion. Thus, the genetic cause of the phenotype in these dogs remains unknown.

In conclusion, we identified a large deletion in the *NSDHL* gene in a female Labrador Retriever and her daughter, which is most likely causative for the hyperkeratotic lesions in these dogs. The identified variant is most likely the result of a recent *de novo* mutation event and not widely distributed in the dog population.

ACKNOWLEDGMENTS

We thank the dog owners who donated samples from their dogs. We acknowledge Ann Hargis for constructive discussions and help in locating biobanked DNA samples of the American Labrador Retrievers with follicular parakeratosis. We thank Nathalie Besuchet Schmutz, Muriel Fragnière, Fabiana Jakob, and Sabrina Schenk for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing the high-throughput sequencing experiments, and the Vital-IT high-performance computing center of the Swiss Institute of Bioinformatics for performing computationally intensive tasks (http://www.vital-it.ch/). We thank the Dog Biomedical Variant Database Consortium (Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Cord Drögemüller, Oliver Forman, Steve Friedenberg, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, V.J., T.L., Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Leonardo Murgiano, Anita Oberbauer, Jeffrey Schoenebeck, Sheila Schmutz, Kim Summers, and Claire Wade) for sharing whole genome sequencing data from control dogs. We also acknowledge all researchers who have deposited dog whole genome sequencing data into public databases.

LITERATURE CITED

Bai, B., W.-M. Zhao, B.-X. Tang, Y.-Q. Wang, L. Wang *et al.*, 2015 DoGSD: the dog and wolf genome SNP database. Nucleic Acids Res. 43: D777–D783.

Bauer, A., D. P. Waluk, A. Galichet, K. Timm, V. Jagannathan *et al.*, 2017 A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. PLoS Genet. 13: e1006651.

Bittar, M., R. Happle, K. H. Grzeschik, L. Leveleki, M. Hertl *et al.*, 2006 CHILD syndrome in 3 generations: the importance of mild or minimal skin lesions. Arch. Dermatol. 142: 348–351.

Bornholdt, D., A. König, R. Happle, L. Leveleki, M. Bittar *et al.*, 2005 Mutational spectrum of NSDHL in CHILD syndrome. J. Med. Genet. 42: e17.

Braverman, N., P. Lin, F. F. Moebius, C. Obie, A. Moser *et al.*, 1999 Mutations in the gene encoding 3 beta-hydroxysteroid-delta 8, delta 7-isomerase cause X-linked dominant Conradi-Hünermann syndrome. Nat Genet. 22: 291–294.

Caldas, H., and G. E. Herman, 2003 NSDHL, an enzyme involved in cholesterol biosynthesis, traffics through the Golgi and accumulates on ER membranes and on the surface of lipid droplets. Hum. Mol. Genet. 12: 2981–2991.

Caldas, H., D. Cunningham, X. Wang, F. Jiang, L. Humphries *et al.*, 2005 Placental defects are associated with male lethality in bare patches and striated embryos deficient in the NAD(P)H steroid dehydrogenase-like (NSDHL) enzyme. Mol. Genet. Metab. 84: 48–60.

Casal, M. L., P. F. Jezyk, J. M. Greek, M. H. Goldschmidt, and D. F. Patterson, 1997 X-linked ectodermal dysplasia in the dog. J. Hered. 88: 513–517.

Cingolani, P., A. Platts, L. Wang, M. Coon, T. Nguyen *et al.*, 2012 A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) 6: 80–92.

Derry, J. M., E. Gormally, G. D. Means, W. Zhao, A. Meindl *et al.*, 1999 Mutations in a delta 8-delta 7 sterol isomerase in the tattered mouse and X-linked dominant chondrodysplasia punctata. Nat. Genet. 22: 286–290.

Grzeschik, K. H., D. Bornholdt, F. Oeffner, A. König, M. del Carmen Boente *et al.*, 2007 Deficiency of PORCN, a regulator of Wnt signaling, is associated with focal dermal hypoplasia. Nat. Genet. 39: 833–835.

Happle, R., 2006 X-chromosome inactivation: role in skin disease expression. Acta Paediatr. Suppl. 95: 16–23.

Happle, R., 2016 The categories of cutaneous mosaicism: a proposed classification. Am. J. Med. Genet. A. 170: 452–459.

Happle, R., H. Koch, and W. Lenz, 1980 The CHILD syndrome. Congenital hemidysplasia with ichthyosiform erythroderma and limb defects. Eur. J. Pediatr. 134: 27–33.

Hargis, A. M., S. Myers, K. Gortel, D. Duclos, and J. Randolph-Habecker, 2013 Proliferative, lymphocytic, infundibular mural folliculitis and dermatitis with prominent follicular apoptosis and parakeratotic casts in four Labrador retrievers: preliminary description and response to therapy. Vet. Dermatol. 24: 346–354, e76–e77.

Heiss, N. S., S. W. Knight, T. J. Vulliamy, S. M. Klauck, S. Wiemann *et al.*, 1998 X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nat. Genet. 19: 32–38.

Kaler, S. G., L. K. Gallo, V. K. Proud, A. K. Percy, Y. Mark *et al.*, 1994 Occipital horn syndrome and a mild Menkes phenotype associated with splice site mutations at the MNK locus. Nat. Genet. 8: 195–202.

Kim, C. A., A. König, D. R. Bertola, L. M. Albano, G. J. Gattás *et al.*, 2005 CHILD syndrome caused by a deletion of exons 6–8 of the *NSDHL* gene. Dermatology 211: 155–158.

König, A., R. Happle, D. Bornholdt, H. Engel, and K. H. Grzeschik, 2000 Mutations in the NSDHL gene, encoding a 3beta-hydroxysteroid dehydrogenase, cause CHILD syndrome. Am. J. Med. Genet. 90: 339–346.

König, A., R. Happle, R. Fink-Puches, H. P. Soyer, D. Bornholdt *et al.*, 2002 A novel missense mutation of NSDHL in an unusual case of CHILD syndrome showing bilateral, almost symmetric involvement. J. Am. Acad. Dermatol. 46: 594–596.

Lemke, J. R., K. Kernland-Lang, K. Hörtnagel, and P. Itin, 2014 Monogenic human skin disorders. Dermatology 229: 55–64.

Lewis, D. T., L. M. Messinger, P. E. Ginn, and M. J. Ford, 1998 A hereditary disorder of cornification and multiple congenital defects in five Rottweiler dogs. Vet. Dermatol. 9: 61–72.

Li, H., 2011 A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987–2993.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760.

Liu, X. Y., A. W. Dangel, R. I. Kelley, W. Zhao, P. Denny *et al.*, 1999 The gene mutated in bare patches and striated mice encodes a novel 3beta-hydroxysteroid dehydrogenase. Nat. Genet. 22: 182–187.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis *et al.*, 2010 The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297–1303.

Murgiano, L., V. Shirokova, M. M. Welle, V. Jagannathan, P. Plattet *et al.*, 2015 Hairless streaks in cattle implicate TSR2 in early hair follicle formation. PLoS Genet. 11: e1005427.

Murgiano, L., D. P. Waluk, R. Towers, N. Wiedemar, J. Dietrich *et al.*, 2016 An intronic MBTPS2 variant results in a splicing defect in horses with brindle coat texture. G3 6: 2963–2970.

Oeffner, F., G. Fischer, R. Happle, A. König, R. C. Betz *et al.*, 2009 IFAP syndrome is caused by deficiency in MBTPS2, an intramembrane zinc metalloprotease essential for cholesterol homeostasis and ER stress response. Am. J. Hum. Genet. 84: 459–467.

Robinson, J. T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander *et al.*, 2011 Integrative genomics viewer. Nat. Biotechnol. 29: 24–26.

Scott, D. W., and W. H. Miller, 2000 Congenital follicular parakeratosis in a Rottweiler and a Siberian Husky. Canine Pract. 25: 16–19.

Smahi, A., G. Courtois, P. Vabres, S. Yamaoka, S. Heuertz *et al.*, 2000 Genomic rearrangement in *NEMO* impairs NF-κB activation and is a cause of incontinentia pigmenti. Nature 405: 466–472.

Towers, R. E., L. Murgiano, D. S. Millar, E. Glen, A. Topf *et al.*, 2013 A nonsense mutation in the *IKBKG* gene in mares with incontinentia pigmenti. PLoS One 8: e81625.

Wang, X., V. Reid Sutton, J. Omar Peraza-Llanes, Z. Yu, R. Rosetta *et al.*, 2007 Mutations in X-linked PORCN, a putative regulator of Wnt signaling, cause focal dermal hypoplasia. Nat. Genet. 39: 836–838.

Communicating editor: D. Bannasch

A frameshift variant in the *EDA* gene in Dachshunds with X-linked hypohidrotic ectodermal dysplasia

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	design, supervision and partial implementation of genetic experiments, review and editing of manuscript

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12729

A frameshift variant in the *EDA* gene in Dachshunds with X-linked hypohidrotic ectodermal dysplasia

S. Hadji Rasouliha*†, A. Bauer*†, M. Dettwiler†‡, M. M. Welle†‡ and T. Leeb*†

*Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland. †DermFocus, University of Bern, 3001 Bern, Switzerland. ‡Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland.

Summary

X-linked hypohidrotic ectodermal dysplasia (XLHED) is a genetic disease characterized by hypoplasia or absence of hair, teeth and sweat glands. The *EDA* gene, located on the X chromosome, encodes the type II transmembrane protein ectodysplasin A. Variants in the *EDA* gene can lead to XLHED in humans, mice, cattle and dogs. In the present study, we investigated a litter of Dachshund puppies, of which four male puppies showed clinical signs of XLHED. We performed a candidate gene analysis in one affected puppy and several non-affected relatives. This analysis revealed a single base-pair deletion in the coding sequence of the *EDA* gene in the affected puppy (NM_001014770.2:c.842delT). The deletion is predicted to cause a frameshift, NP_001014770.1:p.(Leu281HisfsTer22), leading to a premature stop codon which truncates more than one quarter of the EDA protein. Sanger sequencing results confirmed that this variant was inherited from the dam. Based on knowledge about the functional impact of *EDA* variants in dogs and other species, c.842delT is a convincing candidate causative variant for the observed XLHED in the male puppies.

Keywords *Canis lupus familiaris*, dermatology, development, dog, ectodysplasin A, Sanger sequencing, skin, XHED

Ectodermal dysplasia is a group of genetically heterogeneous diseases characterized by abnormal development of two or more ectodermal structures such as hair, teeth, nails and sweat glands (Pinheiro & Freire-Maia 1994). Hypohidrotic ectodermal dysplasia can be inherited through autosomal recessive, autosomal dominant or X-linked modes of inheritance (Pinheiro & Freire-Maia 1994). X-linked hypohidrotic ectodermal dysplasia (XLHED) is one of the most common forms of ectodermal dysplasia and is caused by variants in the *EDA* gene encoding ectodysplasin A (Kere *et al.* 1996; OMIM #305100). Ectodysplasin A is a type II transmembrane protein with several alternative isoforms, of which the longest isoform (isoform 1) consists of 391 amino acid residues. EDAR and XEDAR are two distinct EDA receptors which bind to isoform 1 and isoform 2 of EDA respectively (Yan *et al.* 2000). Once isoform 1 binds to and activates EDAR, it results in translocation of NFκB to the nucleus, where it activates transcription of several target genes. These target genes are required to initiate and differentiate ectoderm-derived tissues such as hair, teeth, nails and sweat glands (Cui & Schlessinger 2006). Several spontaneous *EDA* variants have been previously reported in human XLHED patients and the *tabby* mouse variant (Kere *et al.* 1996; Srivastava *et al.* 1997).

Several independent *EDA* variants have also been described in cattle with XLHED (OMIA 000543-9913; e.g. Drögemüller *et al.* 2001, 2002). In XLHED-affected dogs derived from a German Shepherd, an *EDA* splice site variant was identified (OMIA 000543-9615; Casal *et al.* 2005). These dogs were bred in a colony and used for clinical trials with recombinant ectodysplasin (Casal *et al.* 2007). XLHED-affected crossbred dogs from Israel were found to express a variant *EDA* transcript lacking exon 2. However, the genomic sequence of exon 2 and the conserved splice sites were normal in these dogs. The causative genetic variant in these dogs thus remains unknown (Waluk *et al.* 2016).

In the present study, a litter of four male and three female Dachshund puppies was investigated. All four male littermates had patches of hairless skin on the head and the back. The affected dogs were euthanized during their first three weeks of age. Two excisional skin biopsies were taken

Address for correspondence

T. Leeb, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland.
E-mail: tosso.leeb@vetsuisse.unibe.ch

Accepted for publication 28 August 2018

from the boundary of hairless skin and a normally appearing skin region of one of the affected puppies. Histological examination of these tissue samples revealed a complete absence of adnexal structures, including hair follicles, sebaceous glands and sweat glands in the hairless region, whereas few hair follicles were present in the adjacent haired skin (Fig. 1). In the second tissue sample, which was taken from haired skin, density of follicular compounds was normal. Secondary follicles were lacking, which is normal for very young dogs.

Due to the phenotypic similarity of the affected puppies with earlier described XLHED cases and the fact that only male dogs were affected, we hypothesized that the affected Dachshund puppies also had XLHED and considered *EDA* the top functional candidate gene for this phenotype. We obtained EDTA blood samples from the female littermates, their dam, a brother of the dam and the maternal granddam. Genomic DNA was extracted from the available affected dog and the unaffected relatives. We designed primer pairs for the amplification of all eight exons of the *EDA* gene (Fig. S1, Table S1). PCR products for each *EDA* exon were amplified from genomic DNA using AmpliTaq Gold 360 Master Mix (ThermoFisher). Amplicons were treated with shrimp alkaline phosphatase and endonuclease I, and sequenced on an ABI 3730 capillary sequencer (ThermoFisher). The Sanger sequencing data were then analyzed using the software SEQUENCHER 5.1 (GeneCodes).

The fragment harbouring exon 7 of the *EDA* gene contained a single base deletion—NM_001014770.2: c.842delT, ChrX:g.54 509 504delT (CanFam 3.1) or NP_001014770.1:p.(Leu281HisfsTer22)—in the hemizygous state in the affected puppy (Fig. 2). This frameshift variant leads to the occurrence of a premature stop codon and the replacement of 106 amino acids (27%) of the predicted open reading frame of the wildtype ectodysplasin A protein by 21 new residues (Fig. S1). As the premature stop codon is located very close to the 3′ end of the penultimate exon, it is difficult to predict whether the variant transcript is likely to be subject to nonsense-mediated decay or other cellular quality control mechanisms (Karousis & Mühlemann 2018).

Genotyping the available family members showed that two female littermates of the affected dog and their dam were heterozygous carriers for the variant. The remaining female littermate, the brother of the dam and the maternal granddam did not carry this variant (Fig. 2). As the maternal granddam did not carry the variant in her blood leukocytes, it is likely that the variant arose due to a *de novo* mutation in the dam or in the germ line of one of the maternal grandparents. Unfortunately, no data are available from the maternal grandsire. Hence, it is unknown whether the grandsire had any phenotypic signs of XLHED. It seems unlikely that a male dog with a fully developed, severe XLHED phenotype would have been used as breeding sire, but minimal disease expression in a mosaic dog might have gone undetected.

As of 2015, 82 different *EDA* variants, comprising deletion, missense, nonsense, frameshift and splice variants, were reported in human XLHED patients (Huang *et al.* 2015). In cattle, at least five independent pathogenic *EDA* variants have been described. We now report the third instance of XLHED in dogs and the second instance for which a likely causative genomic variant in the canine *EDA* has been identified. Our study nicely illustrates how a genetic investigation in veterinary medicine may help to confirm a suspected clinical diagnosis. The results are important for developing sustainable breeding programs and decreasing the prevalence of hereditary diseases. In this study, we identified three female carriers, which were removed from breeding to prevent the disease allele from spreading throughout the breed and to avoid the unintentional breeding of additional affected dogs. Simultaneously, a clear female dog was identified that may be used to maintain the particular breeding line.

In summary, we identified a single nucleotide deletion in *EDA* in a male Dachshund with XLHED. The well-known

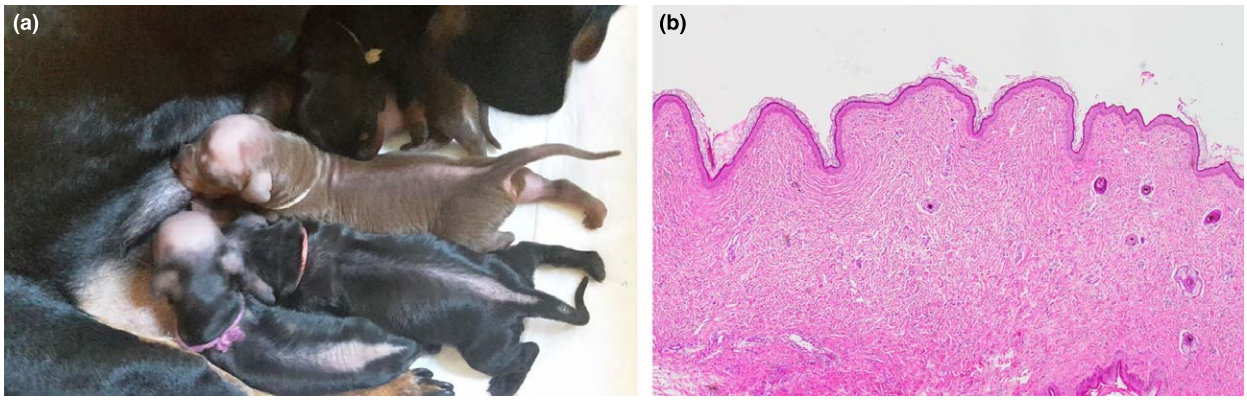


Figure 1 XLHED phenotype in male Dachshund puppies. (a) Patches of hairless skin on the head and back of affected puppies. (b) Histology of a skin sample taken at the boundary of a hairless area (left side, hairless; right side, haired). Adnexal structures (e.g. hair follicles or sebaceous glands) are completely missing on the left side of the sample. Sparse hair follicles are present on the right side. Hematoxylin and Eosin 250×.

 doi: 10.1111/age.12729

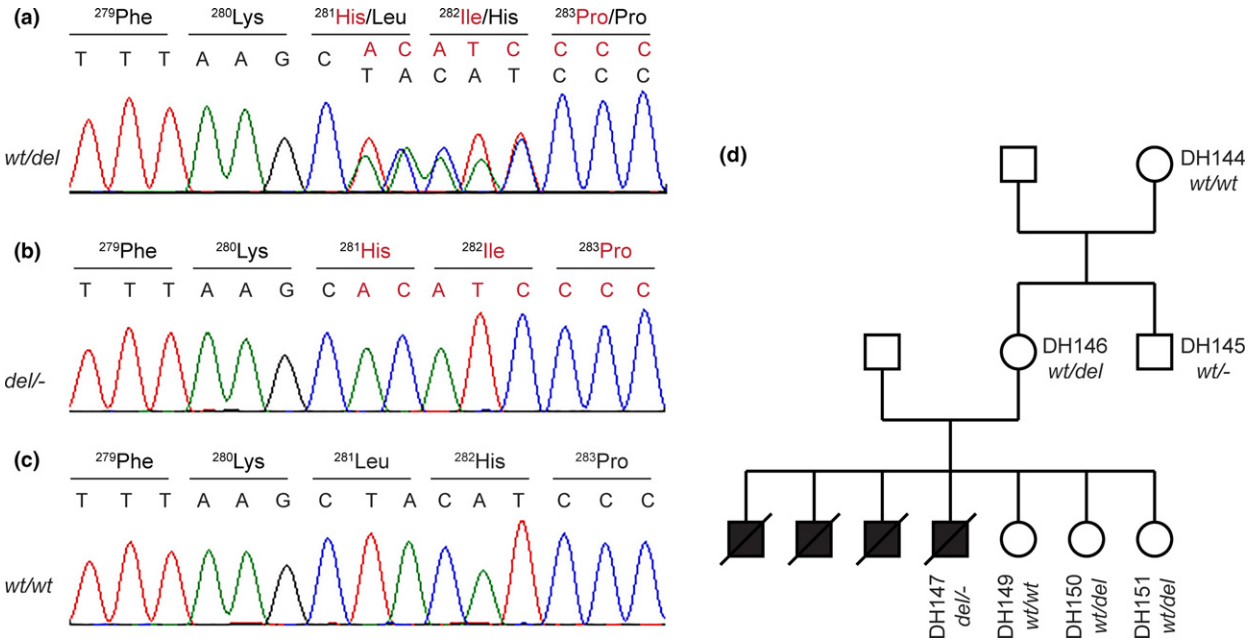


Figure 2 Details of the *EDA*:c.842delT frameshift variant. Wildtype sequences are shown in black, mutant sequences in red. (a) Heterozygous genotype in the dam of the affected puppy. (b) Hemizygous deletion in the affected male puppy. (c) Wildtype sequence in a female littermate of the affected puppy. (d) The pedigree of the family is indicative of an X-linked recessive mode of inheritance and shows co-segregation of the deletion with the phenotype.

functional impact of *EDA* variants in dogs, humans, mice and cattle suggests that c.842delT is a compelling candidate causative variant for the observed XLHED phenotype.

Acknowledgements

The authors would like to thank the dogs' owner for donating samples and pictures and Dr. Melissa Cox from the Center for Animal Genetics (CAG GmbH) for the referral of the case. The authors also wish to thank Nathalie Besuchet Schmutz, Muriel Fragnière and Sabrina Schenk for expert technical assistance. This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738/1) and the Albert Heim Foundation (no. 105).

References

Casal M.L., Scheidt J.L., Rhodes J.L., Henthorn P.S. & Werner P. (2005) Mutation identification in a canine model of X-linked ectodermal dysplasia. *Mammalian Genome* **16**, 524–31.

Casal M.L., Lewis J.R., Mauldin E.A., Tardivel A., Ingold K., Favre M., Paradies F., Demotz S., Gaide O. & Schneider P. (2007) Significant correction of disease after postnatal administration of recombinant ectodysplasin A in canine X-linked ectodermal dysplasia. *American Journal of Human Genetics* **81**, 1050–6.

Cui C. & Schlessinger D. (2006) EDA signaling and skin appendage development. *Cell Cycle* **5**, 2477–83.

Drögemüller C., Distl O. & Leeb T. (2001) Partial deletion of the bovine *ED1* gene causes anhidrotic ectodermal dysplasia in cattle. *Genome Research* **11**, 1699–705.

Drögemüller C., Peters M., Pohlenz J., Distl O. & Leeb T. (2002) A single point mutation within the ED1 gene disrupts correct splicing at two different splice sites and leads to anhidrotic ectodermal dysplasia in cattle. *Journal of Molecular Medicine* **80**, 319–23.

Huang S.X., Liang J.L., Sui W.G., Lin H., Xue W., Chen J.J., Zhang Y., Gong W.W., Dai Y. & Ou M.L. (2015) *EDA* mutation as a cause of hypohidrotic ectodermal dysplasia: a case report and review of the literature. *Genetics and Molecular Research* **14**, 10344–51.

Karousis E.D. & Mühlemann O. (2018) Nonsense-mediated mRNA decay begins where translation ends. *Cold Spring Harbor Perspectives in Biology*. https://doi.org/10.1101/cshperspect.a032862

Kere J., Srivastava A.K., Montonen O. *et al.* (1996) X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. *Nature Genetics* **13**, 409–16.

Pinheiro M. & Freire-Maia N. (1994) Ectodermal dysplasias: a clinical classification and a causal review. *American Journal of Medical Genetics* **53**, 153–62.

Srivastava A.K., Pispa J., Hartung A.J. *et al.* (1997) The Tabby phenotype is caused by mutation in a mouse homologue of the EDA gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 13069–74.

Waluk D., Zur G., Kaufmann R., Welle M.M., Jagannathan V., Drögemüller C., Müller E.J., Leeb T. & Galichet A. (2016) A splice defect in the *EDA* gene in dogs with an X-linked hypohidrotic ectodermal dysplasia (XLHED) phenotype. *G3 (Bethesda)*, **6**, 2949–54.

Yan M., Wang L.C., Hymowitz S.G., Schilbach S., Lee J., Goddard A., de Vos A.M., Gao W.Q. & Dixit V.M. (2000) Two-amino acid molecular switch in an epithelial morphogen that regulates binding to two distinct receptors. *Science* **290**, 523–7.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.
Figure S1 Genomic organization and annotation of the canine *EDA* gene.

Table S1 Primer sequences used for amplifying individual exons and their flanking sequences of the canine *EDA* gene. Coordinates refer to the CanFam3.1 reference genome assembly.

A novel *MLPH* variant in dogs with coat colour dilution

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	genetic analyses, original draft, review and editing of manuscript

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12632

A novel *MLPH* variant in dogs with coat colour dilution

A. Bauer*†, A. Kehl‡, V. Jagannathan*† and T. Leeb*†

*Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland. †DermFocus, University of Bern, 3001 Bern, Switzerland. ‡Laboklin, 97688 Bad Kissingen, Germany.

Summary

Coat colour dilution may be the result of altered melanosome transport in melanocytes. Loss-of-function variants in the *melanophilin* gene (*MLPH*) cause a recessively inherited form of coat colour dilution in many mammalian and avian species including the dog. *MLPH* corresponds to the *D* locus in many domestic animals, and recessive alleles at this locus are frequently denoted with *d*. In this study, we investigated dilute coloured Chow Chows whose coat colour could not be explained by their genotype at the previously known *MLPH*: c.–22G>A variant. Whole genome sequencing of such a dilute Chow Chow revealed another variant in the *MLPH* gene: *MLPH*:c.705G>C. We propose to designate the corresponding mutant alleles at these two variants d^1 and d^2. We performed an association study in a cohort of 15 dilute and 28 non-dilute Chow Chows. The dilute dogs were all either compound heterozygous d^1/d^2 or homozygous d^2/d^2, whereas the non-dilute dogs carried at least one wildtype allele *D*. The d^2 allele did not occur in 417 dogs from diverse other breeds. However, when we genotyped a Sloughi family, in which a dilute coloured puppy had been born out of non-dilute parents, we again observed perfect co-segregation of the newly discovered d^2 allele with coat colour dilution. Finally, we identified a blue Thai Ridgeback with the d^1/d^2 genotype. Thus, our data identify the *MLPH*:c.705G>C as a variant explaining a second canine dilution allele. Although relatively rare overall, this d^2 allele is segregating in at least three dog breeds, Chow Chows, Sloughis and Thai Ridgebacks.

Keywords *Canis lupus familiaris*, melanocyte, melanosome, pigmentation, whole genome sequencing

Coat colour dilution refers to a specific pigmentation phenotype that is found in many mammalian and avian species. This phenotype may be caused by a defect in melanosome transport (Barral & Seabra 2004). Variants in the three genes *RAB27A*, *MYO5A* and *MLPH*, which are indispensable for this process, have been found in humans with Griscelli syndrome types I to III (Pastural *et al.* 1997; Ménasché *et al.* 2000, 2003; Anikster *et al.* 2002). Variants in *RAB27A* and *MYO5A* typically have severe pleiotropic effects and lead to syndromic phenotypes, whereas *MLPH* variants mostly have a more restricted effect on coat colour only. *MLPH* variants are known in many spontaneous mammalian and avian mutants with a dilute coloured phenotype (Matesic *et al.* 2001; Ishida *et al.* 2006; Vaez *et al.* 2008; Bed'hom *et al.* 2012; Cirera *et al.* 2013; Lehner *et al.* 2013; Fontanesi *et al.* 2014; Li *et al.* 2016). In dogs, a recessive non-coding variant at the last nucleotide of exon 1 in the *MLPH* gene (c.–22G>A) is associated with the dilution phenotype in many breeds and used for genetic testing (Philipp *et al.* 2005; Drögemüller *et al.* 2007).

Chow Chow breeders recently were confronted with unexpected genetic testing results, as phenotypically dilute dogs were noticed that were not homozygous for the mutant A allele at *MLPH*:c.–22G>A (Fig. 1). The aim of the present study therefore was to identify the genetic basis of these dilute coloured dogs.

We analysed a cohort of 15 dilute Chow Chows with discordant genetic testing results and 28 non-dilute Chow Chows as controls. Pedigree records were consistent with a monogenic autosomal recessive inheritance of the dilute coat colour phenotype. We initially genotyped all dogs for the known *MLPH*:c.–22G>A variant. For ease of reading, we designate the mutant allele at this variant d^1. None of the 15 dilute Chows was homozygous d^1/d^1: six were heterozygous and nine were homozygous wildtype at this position. We then sequenced the whole genome of a female Chow Chow with blue coat colour (diluted black) that was homozygous *wt*/*wt* at the *MLPH*:c.–22G>A variant. We

Address for correspondence

T. Leeb, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland.
E-mail: Tosso.Leeb@vetsuisse.unibe.ch

Accepted for publication 20 November 2017

Figure 1 Coat colour dilution phenotype and the *MLPH*:c.705G>C variant. Coat colour dilution leads to phenotypically lighter coat colours. Dilute dogs with a black base colour are termed blue, dilute red dogs are termed fawn or cinnamon. (a–d) Representative images of black, blue, red, and fawn Chow Chows respectively are shown. (e) Sloughi puppies at 6 weeks of age. The right-most puppy is dilute, with the black pigment of the mask much lighter and blue eyes (photo courtesy of Ingela Näslund); the dilute coat colour phenotype in this dog became less apparent with aging. (f) Blue Thai Ridgeback. (g) Sanger sequencing electropherograms of dogs with the three genotypes at the *MLPH*:c.705G>C variant. We propose to designate the mutant *C* allele at this variant as d^2.

prepared a PCR-free library, collected 192 197 978 read-pairs of 2 × 150 bp on an Illumina HiSeq 3000 instrument and mapped the reads to the CanFam 3.1 Boxer reference genome yielding a ~23× coverage. The sequence data were submitted to the European Nucleotide Archive (project accession no. PRJEB16012, sample accession no. SAMEA104091566). Single nucleotide variants and short indels were called with respect to the reference genome using GATK (McKenna *et al.* 2010). We filtered for variants that were present in a homozygous alternate state in the

MLPH genotype	*wt/wt*	*wt/d*^{1}	*wt/d*^{2}	*d*^{1}/*d*^{1}	*d*^{1}/*d*^{2}	*d*^{2}/*d*^{2}
Dilute Chow Chows (cases)	—	—	—	—	6	9
Non-dilute Chow Chows (controls)	8	4	16	—	—	—
Dilute Sloughi (case)[1]	—	—	—	—	—	1
Non-dilute Sloughis (controls)[1]	1	—	6	—	—	—
Dilute Thai Ridgeback (case)	—	—	—	—	1	—

[1]The eight Sloughis came from a family consisting of both non-affected parents, five non-affected littermates and one dilute puppy.

Table 1 Association of *MLPH* genotypes with coat colour dilution. The mutant alleles at *MLPH*:c.–22G>A and *MLPH*:c.705G>C are designated *d*^{1} and *d*^{2} respectively.

dilute Chow Chow and heterozygous or homozygous reference in 190 non-dilute control dogs of different breeds and three wolves (Table S1). Among these private variants was a single nucleotide variant at the last nucleotide of exon 7 in the *MLPH* gene. The formally correct variant designation is NM_001103219.2:c.705G>C or Chr25: g.48150787G>C (CanFam3.1 assembly). The variant changes a codon, p.Gln235His. It seems possible that the mutant allele affects splicing, as the wildtype *G* allele corresponds to the consensus of the mammalian 5′-splice site found in ~80% of all mammalian exon/intron junctions, whereas the mutant *C* allele is found in only ~4% of comparable mammalian exon/intron junctions (Sheth *et al.* 2006). As we had no suitable RNA samples, we could not experimentally analyse the effect of the variant on splicing.

We propose to designate the mutant allele at this variant as d^2. We genotyped all available Chow Chows by Sanger sequencing and found a perfect association with the dilution phenotype when considering also compound heterozygosity with the d^1 variant (Table 1, Fig. S1).

Although the vast majority of our control dogs with genome sequence information were homozygous wildtype for the newly discovered *MLPH*:c.705G>C variant, we noticed two Sloughis that were heterozygous. These two Sloughis were parents of a litter that included a puppy with dilute coat colour and blue eyes (Fig. 1e). We genotyped the entire Sloughi family and again found perfect co-segregation of the d^2 allele with the phenotype: Only the blue puppy was homozygous d^2/d^2, whereas the non-dilute littermates were either *wt*/*wt* or *wt*/d^2 (Table 1). Finally, we identified a blue Thai Ridgeback, with the d^1/d^2 genotype (Fig. 1f). The d^2 allele was not present in 417 additional control dogs of 56 different breeds (Table S2).

Given the extensive knowledge about melanophilin function, we think that these data strongly suggest that the newly discovered *MLPH*:c.705G>C variant causes coat colour dilution in dogs and represents another loss-of-function allele similar to *MLPH*:c.–22G>A. This finding of allelic heterogeneity in canine coat colour dilution needs to be considered when performing genetic testing.

Dilute coloured dogs, which are homozygous d^1/d^1, are predisposed for colour dilution alopecia (CDA), a disease characterized by hair loss and chronic inflammation of the skin (Miller 1990; Kim *et al.* 2005; von Bomhard *et al.* 2006; Welle *et al.* 2009). It is currently not fully clear which additional genetic and/or environmental risk factors are involved in CDA. However, CDA seems to be a dog-specific phenomenon. Other known *MLPH*-mutant animals with dilute coloured phenotypes, such as e.g. cats, rabbits, or mice, apparently have no pathological alterations of their hair or skin. Although the coat colour phenotype in d^1/d^1, d^1/d^2 and d^2/d^2 dogs is identical or at least very similar, at this time it is not possible to make a reliable prediction whether these three genotypes show any differences with respect to their risk for CDA. The blue Thai Ridgeback in our study with the d^1/d^2 genotype showed mild signs of CDA. Thus, the association of the *d2* allele with CDA should be carefully investigated before breeding recommendations are given.

Acknowledgements

The authors would like to thank Walter Wurzer, Susan Mothersill, Ingela Näslund, Ullis Sundvell, Mats Guldenheld, Nicole Neukomm and all other involved dog owners for donating samples and pictures and for sharing information of their dogs. The authors also wish to thank Nathalie Besuchet, Muriel Fragnière and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure. We acknowledge collaborators of the Dog Biomedical Variant Database Consortium (DBVDC), Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Cord Drögemüller, Kari Ekenstedt, Kiterie Faller, Oliver Forman, Steve Friedenberg, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Vidhya Jagannathan, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Leonardo Murgiano, Anita Oberbauer, Sheila Schmutz, Jeffrey Schoenebeck, Kim Summers, Frank van Steenbeck, Claire Wade for sharing dog genome sequence data from control dogs and wolves. This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738 / 1) and the Albert-Heim Foundation (no. 105).

Conflicts of interest

Alexandra Kehl is an employee of Laboklin, a company that offers genetic testing as a commercial service.

References

Anikster Y., Huizing M., Anderson P.D., Fitzpatrick D.L., Klar A., Gross-Kieselstein E., Berkun Y., Shazberg G., Gahl W.A. & Hurvitz H. (2002) Evidence that Griscelli syndrome with neurological involvement is caused by mutations in *RAB27A*, not *MYO5A*. *American Journal of Human Genetics* **71**, 407–14.

Barral D.C. & Seabra M.C. (2004) The melanosome as a model to study organelle motility in mammals. *Pigment Cell Research* **17**, 111–8.

Bed'hom B., Vaez M., Coville J.L., Gourichon D., Chastel O., Follett S., Burke T. & Minvielle F. (2012) The lavender plumage colour in Japanese quail is associated with a complex mutation in the region of *MLPH* that is related to differences in growth, feed consumption and body temperature. *BMC Genomics* **13**, 442.

von Bomhard W., Mauldin E.A., Schmutz S.M., Leeb T. & Casal M.L. (2006) Black hair follicular dysplasia in Large Münsterländer dogs: clinical, histological and ultrastructural features. *Veterinary Dermatology* **17**, 182–8.

Cirera S., Markakis M.N., Christensen K. & Anistoroaei R. (2013) New insights into the *melanophilin* (*MLPH*) gene controlling coat color phenotypes in American mink. *Gene* **527**, 48–54.

Drögemüller C., Philipp U., Haase B., Günzel-Apel A.R. & Leeb T. (2007) A noncoding *melanophilin* gene (*MLPH*) SNP at the splice donor of exon 1 represents a candidate causal mutation for coat color dilution in dogs. *Journal of Heredity* **98**, 468–73.

Fontanesi L., Scotti E., Allain D. & Dall'olio S. (2014) A frameshift mutation in the *melanophilin* gene causes the dilute coat colour in rabbit (*Oryctolagus cuniculus*) breeds. *Animal Genetics* **45**, 248–55.

Ishida Y., David V.A., Eizirik E., Schäffer A.A., Neelam B.A., Roelke M.E., Hannah S.S., O'Brien S.J. & Menotti-Raymond M. (2006) A homozygous single-base deletion in *MLPH* causes the dilute coat color phenotype in the domestic cat. *Genomics* **88**, 698–705.

Kim J.H., Kang K.I., Sohn H.J., Woo G.H., Jean Y.H. & Hwang E.K. (2005) Color-dilution alopecia in dogs. *Journal of Veterinary Science* **6**, 259–61.

Lehner S., Gähle M., Dierks C., Stelter R., Gerber J., Brehm R. & Distl O. (2013) Two-exon skipping within *MLPH* is associated with coat color dilution in rabbits. *PLoS One* **8**, e84525.

Li W., Sartelet A., Tamma N., Coppieters W., Georges M. & Charlier C. (2016) Reverse genetic screen for loss-of-function mutations uncovers a frameshifting deletion in the *melanophilin* gene accountable for a distinctive coat color in Belgian Blue cattle. *Animal Genetics* **47**, 110–3.

Matesic L.E., Yip R., Reuss A.E., Swing D.A., O'Sullivan T.N., Fletcher C.F., Copeland N.G. & Jenkins N.A. (2001) Mutations in *Mlph*, encoding a member of the Rab effector family, cause the melanosome transport defects observed in leaden mice. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10238–43.

McKenna A., Hanna M., Banks E. *et al.* (2010) The GENOME ANALYSIS TOOLKIT: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297–303.

Ménasché G., Pastural E., Feldmann J. *et al.* (2000) Mutations in *RAB27A* cause Griscelli syndrome associated with haemophagocytic syndrome. *Nature Genetics* **25**, 173–6.

Ménasché G., Ho C.H., Sanal O., Feldmann J., Tezcan I., Ersoy F., Houdusse A., Fischer A. & de Saint Basile G. (2003) Griscelli syndrome restricted to hypopigmentation results from a *melanophilin* defect (GS3) or a *MYO5A* F-exon deletion (GS1). *Journal of Clinical Investigation* **112**, 450–6.

Miller W.H. (1990) Colour dilution alopecia in Doberman Pinschers with blue or fawn coat colours: A study of the incidence and histopathology of this disorder. *Veterinary Dermatology* **1**, 113–22.

Pastural E., Barrat F.J., Dufourcq-Lagelouse R., Certain S., Sanal O., Jabado N., Seger R., Griscelli C., Fischer A. & de Saint Basile G. (1997) Griscelli disease maps to chromosome 15q21 and is associated with mutations in the myosin-Va gene. *Nature Genetics* **16**, 289–92.

Philipp U., Hamann H., Mecklenburg L., Nishino S., Mignot E., Günzel-Apel A.R., Schmutz S.M. & Leeb T. (2005) Polymorphisms within the canine *MLPH* gene are associated with dilute coat color in dogs. *BMC Genetics* **6**, 34.

Sheth N., Roca X., Hastings M.L., Roeder T., Krainer A.R. & Sachidanandam R. (2006) Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Research* **34**, 3955–67.

Vaez M., Follett S.A., Bed'hom B., Gourichon D., Tixier-Boichard M. & Burke T. (2008) A single point-mutation within the *melanophilin* gene causes the lavender plumage colour dilution phenotype in the chicken. *BMC Genetics* **9**, 7.

Welle M., Philipp U., Rüfenacht S. *et al.* (2009) *MLPH* genotype melanin phenotype correlation in dilute dogs. *Journal of Heredity* **100**, S75–9.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1 *MLPH* haplotypes confirm compound heterozygosity.

Table S1 Accession numbers of 191 dog/wolf genome sequences.

Table S2 Genotypes of 417 dogs from different breeds.

A second *KRT71* allele in curly coated dogs

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	contribution to sample collection, design and implementation of genetic experiments, original draft, review and editing of manuscript, illustrations

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12743

A second *KRT71* allele in curly coated dogs

A. Bauer*†, S. Hadji Rasouliha*†, M. T. Brunner†‡, V. Jagannathan*†, I. Bucher†§, J. Bannoehr†§¶, K. Varjonen††, R. Bond**, K. Bergvall‡‡, M. M. Welle†‡, P. Roosje†§ and T. Leeb*†**

*Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland. †Dermfocus, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland. ‡Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland. §Division of Clinical Dermatology, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern, Bern, 3001, Switzerland. ¶Dermatology Department, Animal Health Trust, Lanwades Park, Kentford, Newmarket, Suffolk, CB8 7UU, UK. **Department Clinical Sciences and Services, Royal Veterinary College, Hatfield, AL9 7TA, UK. ††Anicura Albano Animal Hospital, Danderyd, 18236, Sweden. ‡‡Department of Clinical Sciences, Swedish University of Agricultural Sciences, Box 7054, Uppsala, 750 07, Sweden.

Summary

Major characteristics of coat variation in dogs can be explained by variants in only a few genes. Until now, only one missense variant in the *KRT71* gene, p.Arg151Trp, has been reported to cause curly hair in dogs. However, this variant does not explain the curly coat in all breeds as the mutant ^{151}Trp allele, for example, is absent in Curly Coated Retrievers. We sequenced the genome of a Curly Coated Retriever at 22× coverage and searched for variants in the *KRT71* gene. Only one protein-changing variant was present in a homozygous state in the Curly Coated Retriever and absent or present in a heterozygous state in 221 control dogs from different dog breeds. This variant, NM_001197029.1: c.1266_1273delinsACA, was an indel variant in exon 7 that caused a frameshift and an altered and probably extended C-terminus of the KRT71 protein NP_001183958.1: p.(Ser422ArgfsTer?). Using Sanger sequencing, we found that the variant was fixed in a cohort of 125 Curly Coated Retrievers and segregating in five of 14 additionally tested breeds with a curly or wavy coat. *KRT71* variants cause curly hair in humans, mice, rats, cats and dogs. Specific *KRT71* variants were further shown to cause alopecia. Based on this knowledge from other species and the predicted molecular consequence of the newly identified canine *KRT71* variant, it is a compelling candidate causing a second curly hair allele in dogs. It might cause a slightly different coat phenotype than the previously published p.Arg151Trp variant and could potentially be associated with follicular dysplasia in dogs.

Keywords animal model, *canis lupus familiaris*, dermatology, hair, hair follicle, keratin, morphology, whole genome sequencing

In 2009, Cadieu and colleagues reported that major characteristics of coat variation in dogs are influenced by variants in three genes. A missense variant in *FGF5* explained long hair, a 167-bp insertion into the 3′-UTR of *RSPO2* was reported to cause wiry hair and furnishings (increased hair growth on face and legs) and a *KRT71* missense variant (c.451C>T, p.Arg151Trp) was found in dogs with a curly or wavy coat (Cadieu *et al.* 2009). Specific combinations of alleles with these variants lead to the desired breed standards, and genetic testing enables breeders to avoid unwanted coat phenotypes to a certain degree. However, in some cases genetic testing for the known variants fails to explain the coat texture. According to dog owners, the commercially testable *KRT71*:c.451C>T variant does not explain the curly hair in Curly Coated Retrievers. The standard coat of a Curly Coated Retriever is a black or liver colour with tight, crisp curls on the body and smooth, straight and very short hair on face and legs. Curly Coated Retrievers often have a sparsely haired tail ('rat tail'). Follicular dysplasia, characterized by symmetrical, non-pruritic hair loss, often with a waxing and waning course, is a relatively common coat disorder in this breed (Bond *et al.* 2016). During breed formation, Curly Coated Retrievers most likely got their curly coat from Irish Water Spaniels, another breed that is predisposed to follicular dysplasia (Cerundolo *et al.* 1999, 2000; Bond *et al.* 2016). A genetic background for this disorder is suspected. In the present study, we aimed to identify the causative variant for the curly coat in Curly Coated Retrievers.

Address for correspondence

T. Leeb, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, Bern 3001, Switzerland.
E-mail: Tosso.Leeb@vetsuisse.unibe.ch

Accepted for publication 09 October 2018

Given that the known *KRT71*:c.451C>T variant was not present in several tested Curly Coated Retrievers, we hypothesized that another genetic variant was likely to have caused their curls and is probably fixed in the breed. We therefore sequenced the genome from a Curly Coated Retriever at 22× coverage (ENA project accession no. PRJEB16012, sample accession, SAMEA104091556) and compared the sequencing data to 221 dog genomes from different breeds (Table S1). We called single nucleotide variants and short indels with respect to the canine reference genome assembly CanFam 3.1 as described earlier (Bauer *et al.* 2017). Applying a candidate gene approach, we extracted variants in the *KRT71* gene that were present in any of the 222 canine whole genome sequences (Table S2). Focusing on protein-changing variants, we found a single *KRT71* indel variant present in a homozygous state in the Curly Coated Retriever that was absent or heterozygous in all other dogs. The variant can be formally described as NM_001197029.1:c.1266_1273delinsACA or NP_001183958.1:p.(Ser422ArgfsTer?). The identified novel *KRT71* variant is predicted to cause a frameshift with an extended mutant reading frame compared to the wildtype transcript, KRT71:p.(Ser422ArgfsTer212). Compared to the wildtype 525-amino-acid KRT71, the predicted mutant protein has a length of 632 amino acids (Fig. S1). RNA-seq data from skin biopsies demonstrated expression of the expected transcript without any splicing alterations from the allele with the indel variant (Fig. 1; Table S3).

To investigate the *KRT71* allele and genotype distribution in different dog breeds with curly or wavy coat, we genotyped 1286 dogs from 15 breeds for the previously described c.451C>T missense variant and the c.1266_1273delinsACA variant using Sanger sequencing (Table 1). For ease of reading we will from now on refer to the mutant alleles at these two variants as c^1 and c^2, respectively.

All 125 tested Curly Coated Retrievers were homozygous wildtype at c.451C>T and homozygous mutant c^2/c^2 at c.1266_1273delinsACA. Although the c^1 allele overall was more common than the c^2 allele, there were many curly coated breeds in which both alleles segregated. We also observed two dogs carrying both variants on one haplotype, suggesting a past recombination or gene conversion event.

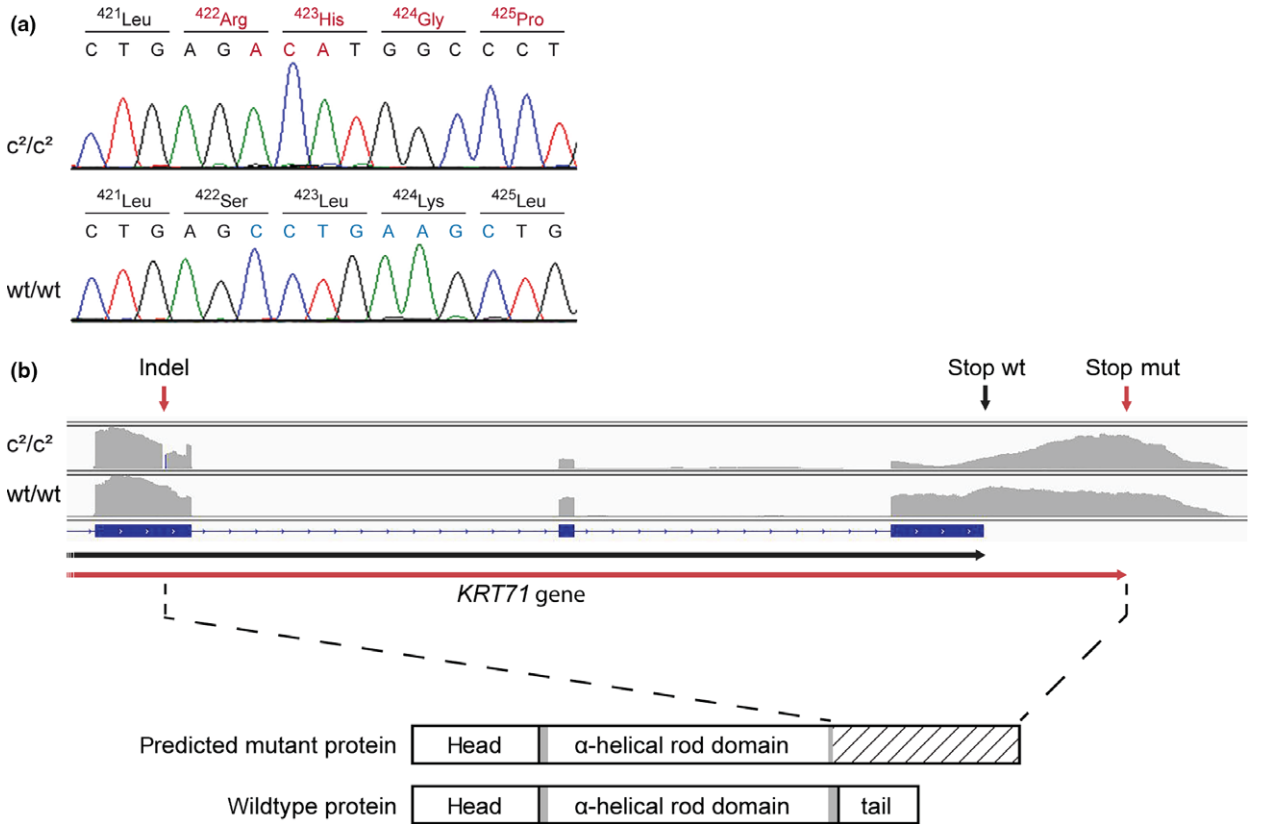


Figure 1 Illustration of the *KRT71*:c.1266_1273delinsACA variant and its predicted consequence on the KRT71 protein. (a) Sanger electropherograms of dogs with c^2/c^2 and *wt/wt* genotypes. Bases shown in red are inserted, bases in blue are deleted in the mutant c^2 allele. The variant leads to a shift in the reading frame and thus an altered amino acid sequence in the mutant protein (amino acids written in red letters). (b) RNA-seq data from a Curly Coated Retriever with c^2/c^2 genotype and a Greyhound with *wt/wt* genotype showing the read coverage mapped to the last three exons of the *KRT71* gene. As a result of the indel variant in exon 7 (red arrow), the mutant open reading frame is extended and ends at a later termination codon (stop mut, red arrow) compared to the wildtype open reading frame (ending at stop *wt*). A schematic of the predicted mutant and wildtype protein is shown below the RNA-seq data. Grey boxes indicate helix initiation and helix termination motifs. The frameshift is predicted to affect part of the helix termination motif and the whole C-Terminal tail of the protein (hatched area).

Table 1 *KRT71* diplotypes in 1286 dogs from 15 different breeds with curly or wavy hair. We refer to the mutant alleles at the previously described c.451C>T missense variant as c^1 and the c.1266_1273delinsACA variant as c^2. In five of the tested breeds, both alleles were segregating.

	KRT71 diplotype							
Breed	*wt/wt*	c^1/wt	c^1/c^1	c^2/wt	c^1/c^2	c^2/c^2	c^1/c^1c^2	*n*
Airedale Terrier			5					5
Barbet	44	2						46
Bergamasco Shepherd dog	8	4						12
Bolonka Zwetna	8	8	8					24
Cão de Serra de Aires	15							15
Chesapeake Bay Retriever			3		9	1[1]		13
Curly Coated Retriever						125		125
Lagotto Romagnolo		14	559	3	172	12[1]		760
Mudi			11		18	3	2	34
Perro de Agua Español			73		20	2[1]		95
Poodle			89		6	1[2]		96
Portugese Water dog		3						3
Puli			1					1
Schapendoes	44	4						48
Soft Coated Wheaten Terrier	7	2						9
Total								1286

[1]The cohort comprised three dogs that had been diagnosed with follicular dysplasia. They all had the diplotype c^2/c^2.
[2]This poodle had pronounced alopecia on the body and an almost completely hairless tail. Unfortunately, no histopathological examination of a skin biopsy was performed to confirm the suspected follicular dysplasia.

This rare haplotype should be considered during future genetic testing applications.

KRT71 encodes a type II keratin specifically expressed in the inner root sheath of the hair follicle (Aoki *et al.* 2001; Langbein *et al.* 2002). Other than in dogs, *KRT71* variants have been identified and studied at a molecular level in humans, mice, rats and cats. In humans, a *KRT71* missense variant affecting the helix initiation motif was discovered in a Japanese family with autosomal dominant woolly hair and hypotrichosis (Fujimoto *et al.* 2012). In mice, different dominant *Krt71* variants were described in spontaneous or induced mutants with curly or wavy hair such as the caracul or the Rco12 and Rco13 mutants (Kikkawa *et al.* 2003; Runkel *et al.* 2006). Furthermore, a recessive 10-bp deletion variant in exon 1 of *Krt71* was reported to cause progressive alopecia in the Rco3 mouse mutant (Peters *et al.* 2003). In rats, a *Krt71* splice defect was identified as the semidominant *Rex* allele that causes curly hair in heterozygous animals and hair loss in homozygous mutant rats (Kuramoto *et al.* 2010). In rexoid cats, *KRT71* splice defects were found in the Devon Rex breed and in curly coated Selkirk Rex cats (Gandolfi *et al.* 2010, 2013). Finally, in another cat breed, the hairless Sphinx, a splice site variant in *KRT71* leading to a frameshift and a premature stop codon was found, mostly in a homozygous state or compound heterozygous with the Devon Rex allele (Gandolfi *et al.* 2010).

The canine c.1266_1273delinsACA variant identified in the present study affects the sequence coding for the helix termination motif, altering approximately three-quarters of its sequence as well as the whole C-terminal tail. In view of the numerous examples of *KRT71* variants causing curly hair, the inner root sheath restricted expression of KRT71 in the hair follicle and the predicted effect of the variant c.1266_1273delinsACA on the protein, we think that it is a compelling candidate causative variant for the curly coat in Curly Coated Retrievers and other dog breeds. As specific *KRT71* variants can lead to hair loss in addition to curls in humans, mice, rats and cats, we further hypothesized that the canine c^2 allele might represent a genetic risk factor for some forms of follicular dysplasia in dogs. Because Curly Coated Retrievers are fixed for the c^2 allele, we could not test for an association with follicular dysplasia in this breed. However, it is noteworthy that the Curly Coated Retriever breed standard calls for a smooth and straight coat on the forehead, face, front of forelegs and feet (http://images.akc.org/pdf/breeds/standards/CurlyCoatedRetriever.pdf). This is quite distinct from breeds such as the Poodle, which has curls on the entire coat. Poodles are very nearly fixed for the c^1 allele. We thus speculate that the functional effect of the c^2 allele may be slightly different from that of the c^1 allele and that this hypothetical functional difference may be involved in the expression of the characteristic coat phenotype of Curly Coated Retrievers.

Our cohort contained three curly coated dogs (one Chesapeake Bay Retriever, one Lagotto Romagnolo, one Perro de Agua Español) affected with follicular dysplasia that had been diagnosed by veterinary pathologists. All three follicular dysplasia cases were homozygous c^2/c^2. Our cohort also contained five Lagotto Romagnolos from Sweden diagnosed with adult onset spontaneous, symmetrical, non-inflammatory alopecia affecting the trunk that were homozygous c^1/c^1. They were otherwise healthy.

In the Lagotto Romagnolo breed, c^1 and c^2 alleles were observed in all possible genotypic combinations. Initially, we

 doi: 10.1111/age.12743

did not have phenotype information on the coat quality of most sampled Lagotto Romagnolo dogs in our biobank. Upon follow up of the c^2/c^2 dogs, one of the owners reported severe alopecia on the back (Fig. S2). Other c^2/c^2 Lagotto Romagnolo dogs showed a normal coat, but several of their owners reported that the hair of these dogs could be plucked very easily.

Given the low frequency of the c^2 allele in Poodles, it is very striking that the only identified homozygous c^2/c^2 Poodle was alopecic and the potential differentials included follicular dysplasia. Similarly, the only observed homozygous c^2/c^2 Chesapeake Bay Retriever was diagnosed with follicular dysplasia. On the other hand, across breeds only a relatively small proportion of homozygous c^2/c^2 dogs in our study showed hair loss. It is therefore conceivable that the c^2 allele possibly represents a predisposing genetic risk factor for follicular dysplasia. However, even if this hypothesis is correct, other genetic and/or environmental factors will be required for the manifestation of follicular dysplasia.

In conclusion, we identified the indel variant *KRT71*:c.1266_1273delinsACA indel variant as the most likely causal variant underlying a second curly hair allele in dogs. This allele possibly causes a slightly different coat phenotype as the previously described curly allele with the *KRT71*:c.451C>T missense variant. Further studies are required to confirm whether the new curly allele modulates the risk for follicular dysplasia.

Acknowledgements

The authors would like to thank the dog owners for donating samples and pictures and for sharing information about their dogs. The authors also wish to thank Nathalie Besuchet, Muriel Fragnière and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure. This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738/1) and the Albert-Heim Foundation (no. 105).

References

Aoki N., Sawada S., Rogers M.A., Schweizer J., Shimomura Y., Tsujimoto T., Ito K. & Ito M. (2001) A novel type II cytokeratin, mK6irs, is expressed in the Huxley and Henle layers of the mouse inner root sheath. *Journal of Investigative Dermatology* **116**, 359–65.

Bauer A., Waluk D.P., Galichet A. *et al.* (2017) A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. *PLoS Genetics* **13**, e1006651.

Bond R., Varjonen K., Hendricks A., Chang Y.M. & Brooks Brownlie H. (2016) Clinical and pathological features of hair coat abnormalities in curly coated retrievers from UK and Sweden. *Journal of Small Animal Practice* **57**, 659–67.

Cadieu E., Neff M.W., Quignon P. *et al.* (2009) Coat variation in the domestic dog is governed by variants in three genes. *Science* **326**, 150–3.

Cerundolo R., Lloyd D.H. & Pidduck H.G. (1999) Studies on the inheritance of hair loss in the Irish water spaniel. *Veterinary Record* **145**, 542–4.

Cerundolo R., Lloyd D.H., McNeil P.E. & Evans H. (2000) An analysis of factors underlying hypotrichosis and alopecia in Irish Water Spaniels in the United Kingdom. *Veterinary Dermatology* **11**, 107–22.

Fujimoto A., Farooq M., Fujikawa H. *et al.* (2012) A missense mutation within the helix initiation motif of the keratin *K71* gene underlies autosomal dominant woolly hair/hypotrichosis. *Journal of Investigative Dermatology* **132**, 2342–9.

Gandolfi B., Outerbridge C.A., Beresford L.G., Myers J.A., Pimentel M., Alhaddad H., Grahn J.C., Grahn R.A. & Lyons L.A. (2010) The naked truth: Sphynx and Devon Rex cat breed mutations in *KRT71*. *Mamalian Genome* **21**, 509–15.

Gandolfi B., Alhaddad H., Joslin S.E., Khan R., Filler S., Brem G. & Lyons L.A. (2013) A splice variant in *KRT71* is associated with curly coat phenotype of Selkirk Rex cats. *Scientific Reports* **3**, 2000.

Kikkawa Y., Oyama A., Ishii R. *et al.* (2003) A small deletion hotspot in the type II keratin gene *mK6irs1*/*Krt2-6g* on mouse chromosome 15, a candidate for causing the wavy hair of the caracul (Ca) mutation. *Genetics* **165**, 721–33.

Kuramoto T., Hirano R., Kuwamura M. & Serikawa T. (2010) Identification of the rat Rex mutation as a 7-bp deletion at splicing acceptor site of the *Krt71* gene. *Journal of Veterinary Medical Science* **72**, 909–12.

Langbein L., Rogers M.A., Praetzel S., Aoki N., Winter H. & Schweizer J. (2002) A novel epithelial keratin, hK6irs1, is expressed differentially in all layers of the inner root sheath, including specialized huxley cells (Flugelzellen) of the human hair follicle. *Journal of Investigative Dermatology* **118**, 789–99.

Peters T., Sedlmeier R., Büssow H. *et al.* (2003) Alopecia in a novel mouse model RCO3 is caused by mK6irs1 deficiency. *Journal of Investigative Dermatology* **121**, 674–80.

Runkel F., Klaften M., Koch K., Bohnert V., Bussow H., Fuchs H., Franz T. & Hrabe de Angelis M. (2006) Morphologic and molecular characterization of two novel *Krt71* (Krt2-6 g) mutations: Krt71rco12 and Krt71rco13. *Mammalian Genome* **17**, 1172–82.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Alignments of wildtype and mutant canine *KRT71* transcripts and KRT71 proteins.

Figure S2 Coat phenotypes in dogs with different *KRT71* genotypes.

Table S1 Sample designations and breed information on 222 dogs with genome sequences.

Table S2 Genotypes of 222 dogs at *KRT71* variants where the sequenced Curly Coated Retriever was homozygous for the mutant allele.

Table S3 Sample designations and breed information on dogs with RNA-seq data.

A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis

Journal:	PLoS Genetics
Manuscript status:	published
Contributions:	genetic analyses, original draft, review and editing of manuscript, illustrations

PLOS | GENETICS

RESEARCH ARTICLE

A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis

Anina Bauer[1,2], Dominik P. Waluk[2,3], Arnaud Galichet[2,3], Katrin Timm[2,4], Vidhya Jagannathan[1,2], Beyza S. Sayar[2,3], Dominique J. Wiener[2,5,6], Elisabeth Dietschi[1], Eliane J. Müller[2,3,5,7], Petra Roosje[2,8], Monika M. Welle[2,5], Tosso Leeb[1,2]*

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **2** DermFocus, University of Bern, Bern, Switzerland, **3** Department of Clinical Research, Molecular Dermatology and Stem Cell Research, University of Bern, Bern, Switzerland, **4** Dermavet, Oberentfelden, Switzerland, **5** Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **6** Hubrecht Institute, CT Utrecht, The Netherlands, **7** Clinic for Dermatology, Inselspital, Bern University Hospital, Bern, Switzerland, **8** Division of Clinical Dermatology, Vetsuisse Faculty, University of Bern, Bern, Switzerland

* Tosso.Leeb@vetsuisse.unibe.ch

OPEN ACCESS

Citation: Bauer A, Waluk DP, Galichet A, Timm K, Jagannathan V, Sayar BS, et al. (2017) A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. PLoS Genet 13(3): e1006651. doi:10.1371/journal.pgen.1006651

Editor: Judith Fischer, University Medical Center Freiburg, GERMANY

Received: October 24, 2016

Accepted: February 23, 2017

Published: March 1, 2017

Data Availability Statement: The sequence data were deposited under the study accession PRJEB16012 at the European Nucleotide Archive. The sample accessions are SAMEA4506895 for the case (DS043), SAMEA72802918 for the sire (DS053) and SAMEA72802168 for the dam (DS051).

Funding: This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738 / 1) and the Albert-Heim Foundation. The funders had no role in study

Abstract

Ichthyoses are a heterogeneous group of inherited cornification disorders characterized by generalized dry skin, scaling and/or hyperkeratosis. Ichthyosis vulgaris is the most common form of ichthyosis in humans and caused by genetic variants in the *FLG* gene encoding filaggrin. Filaggrin is a key player in the formation of the stratum corneum, the uppermost layer of the epidermis and therefore crucial for barrier function. During terminal differentiation of keratinocytes, the precursor profilaggrin is cleaved by several proteases into filaggrin monomers and eventually processed into free amino acids contributing to the hydration of the cornified layer. We studied a German Shepherd dog with a novel form of ichthyosis. Comparing the genome sequence of the affected dog with 288 genomes from genetically diverse non-affected dogs we identified a private heterozygous variant in the *ASPRV1* gene encoding "aspartic peptidase, retroviral-like 1", which is also known as skin aspartic protease (SASPase). The variant was absent in both parents and therefore due to a *de novo* mutation event. It was a missense variant, c.1052T>C, affecting a conserved residue close to an autoprocessing cleavage site, p.(Leu351Pro). *ASPRV1* encodes a retroviral-like protease involved in profilaggrin-to-filaggrin processing. By immunofluorescence staining we showed that the filaggrin expression pattern was altered in the affected dog. Thus, our findings provide strong evidence that the identified *de novo* variant is causative for the ichthyosis in the affected dog and that ASPRV1 plays an essential role in skin barrier formation. *ASPRV1* is thus a novel candidate gene for unexplained human forms of ichthyoses.

Author summary

The skin undergoes a constant process of self-renewing and keratinocytes migrate from the basal layer of the epidermis to the uppermost layer, the stratum corneum, as they differentiate. A defect in the differentiation of keratinocytes can lead to cornification

design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

disorders such as ichthyosis. The most common form of this disorder in humans is ichthyosis vulgaris caused by variants in the filaggrin gene. Filaggrin is required for bundling intermediate filaments resulting in the flattening of keratinocytes. Filaggrin is produced from profilaggrin and the processing steps involve several enzymes including proteases. In the present study, we sequenced the genome of a dog with a novel form of ichthyosis. By comparing this sequence to 288 control genomes, we identified a private missense variant in the *ASPRV1* gene encoding the retroviral-like aspartic protease 1, also known as SASPase, which is involved in the processing of profilaggrin. The variant was due to a *de novo* mutation event, which is consistent with the patient being an isolated single case of a novel form of ichthyosis. Filaggrin protein expression was altered in the skin of the affected dog. Thus, our results strongly suggest that genetic variants in *ASPRV1* can cause ichthyosis by altering filaggrin processing.

Introduction

The skin and in particular the epidermis provide both an outward and inward barrier function, which is essential for survival. Aberrant skin development or homeostasis can impair this barrier function and may result in skin disorders. Ichthyoses are a heterogeneous group of skin disorders characterized by dry skin, scaling and/or hyperkeratosis, often associated with erythroderma [1,2].These clinical signs are caused by a defect in the terminal differentiation of keratinocytes and subsequent desquamation taking place in the uppermost layer of the epidermis, the stratum corneum. Ichthyoses are primarily inherited skin disorders that can either be non-syndromic, when clinical findings are limited to the skin, or syndromic in case additional organs are involved [2]. Non-syndromic forms of ichthyoses are further sub-classified into common ichthyoses, autosomal recessive congenital ichthyoses, keratinopathic ichthyoses caused by variants in different keratin genes, and other forms of ichthyoses [1,2].

The common ichthyoses consist of ichthyosis vulgaris (IV) and recessive X-linked ichthyosis (RXLI). IV is the most common and mildest form of ichthyosis with an incidence of approximately 1:250 to 1:1000 in humans [2,3]. IV is caused by different semidominant genetic variants in the *FLG* gene encoding filaggrin [4]. Filaggrin (filament aggregating protein) is a keratin bundling protein and a key player in the formation of the stratum corneum [5]. The precursor of filaggrin is the >400 kDa protein profilaggrin, which is the major component of keratohyalin granules in the granular layer of the skin [6,7]. Profilaggrin consists of a unique N-terminus and a series of filaggrin units separated by short linker peptides. The initially highly phosphorylated profilaggrin is dephosphorylated and cleaved by proteases into individual filaggrin molecules during the cornification process. In addition to its role in aggregating keratin intermediate filaments into bundles, filaggrin is also degraded into free amino acids that contribute to the hydration of the cornified layer [8].

RXLI, sometimes also called X-linked ichthyosis (XLI), is clinically more severe and characterized by dark brown scales and generalized dry skin. It is caused by variants, mainly large deletions, affecting the *STS* gene, which encodes steroid sulfatase [9].

Autosomal recessive congenital ichthyoses (ARCI) are the second category of non-syndromic ichthyoses. They may be caused by variants in at least 9 different genes: *ABCA12*, *ALOX12B*, *ALOXE3*, *CERS3*, *CYP4F22*, *LIPN*, *NIPAL4*, *PNPLA1* and *TGM1* [2].

Finally, keratinopathic ichthyoses, the third category of non-syndromic ichthyoses, are caused by variants in the *KRT1*, *KRT2*, or *KRT10* genes [2]. Thus, there are currently 14 human genes implicated in different forms of non-syndromic ichthyoses [2,10].

Dogs represent valuable models for many human hereditary diseases and enabled for example the discovery of *PNPLA1* as an ichthyosis gene. The first pathogenic *PNPLA1* variant was identified in Golden Retriever ichthyosis, which is characterized by a mild phenotype. Interestingly, this canine genodermatosis currently has an extremely high prevalence in the breed [11,12]. Other dog models for human non-syndromic ichthyoses include Norfolk Terriers with an epidermolytic ichthyosis caused by a *KRT10* variant [13], Bulldogs with ARCI caused by a *NIPAL4* variant [14], and Jack Russell Terriers with another form of ARCI caused by a *TGM1* variant [15]. Further cases of canine ichthyoses have been reported, but the underlying genetic defects have not been solved [16]. Thus, dogs might help to identify additional ichthyosis genes, which might be of relevance for unsolved human forms of ichthyoses.

In the present study, we describe a novel non-epidermolytic form of ichthyosis in a German Shepherd. In this breed, until now, no ichthyosis cases have been reported in the scientific literature. We therefore applied a whole genome sequencing approach to unravel the causative genetic variant.

Results

Clinical and histopathological phenotype

An intact female German Shepherd was presented at 10 months of age with a history of severe scaling of the skin with mild pruritus. According to the owner, the lesions started to develop shortly after birth. Dermatological examination revealed generalized hypotrichosis and focal areas of alopecia with generalized severe exfoliation of greyish scales and mild erythema. Comedones were seen on the ventral abdomen and in the perivulvar area (Fig 1).

The owner reported that this phenotype had not been seen in the six littermates or the parents of the affected German Shepherd. The skin condition improved under topical treatment with a rehydrating, anti-seborrheic spray and shampoo.

Histopathological analysis of four skin biopsies from different body regions revealed a severe laminar to compact orthokeratotic hyperkeratosis extending into the follicular

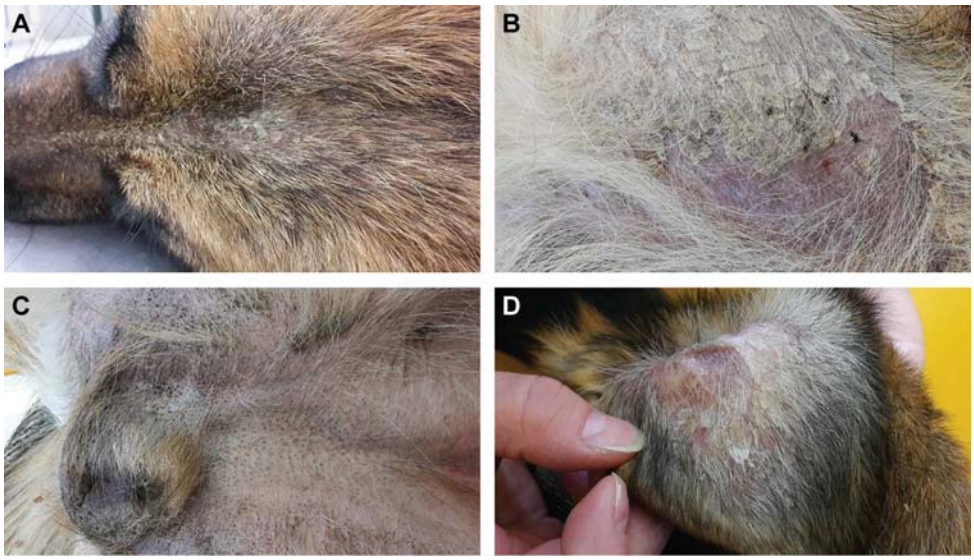


Fig 1. Clinical phenotype of the affected German Shepherd. (A) Hypotrichosis and scaly skin. **(B)** Alopecic region of the thigh with erythema and scales. **(C)** Comedones in the inguinal region. **(D)** Pinna with scales and erythema.

doi:10.1371/journal.pgen.1006651.g001

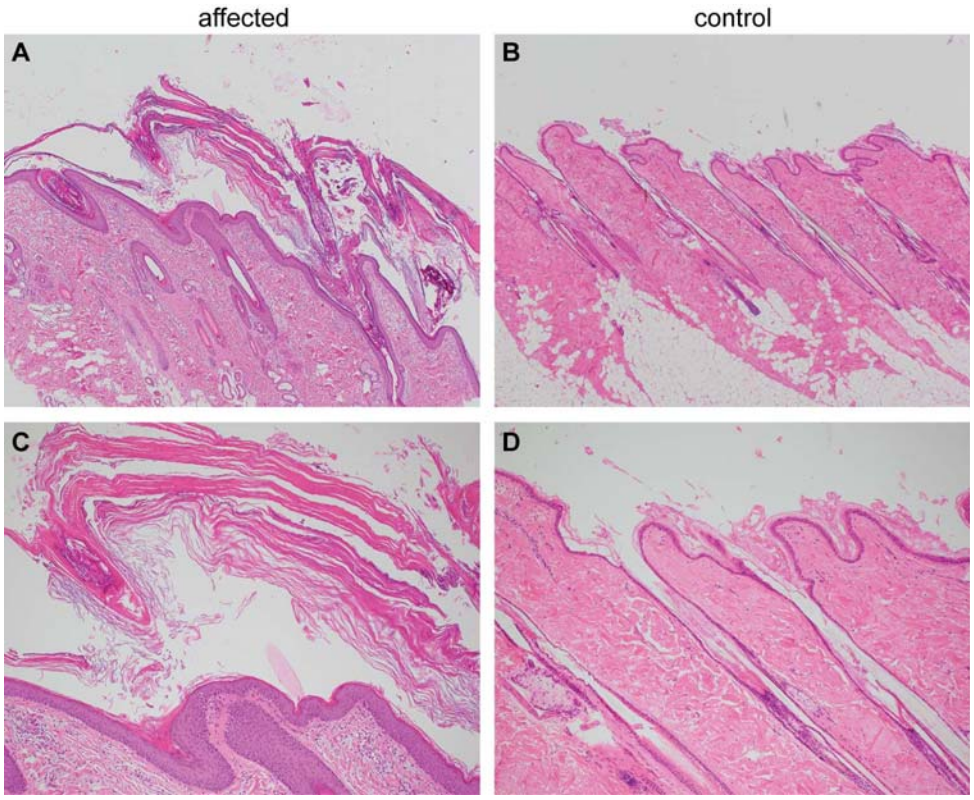


Fig 2. Histopathological findings in skin of the ichthyotic dog and a control dog. (A) Skin of the ichthyotic dog with severe laminar to compact orthokeratotic hyperkeratosis extending in the follicular infundibula and covering a mildly hyperplastic epidermis. Hematoxylin and Eosin 40x. **(B)** Skin of a normal dog with normal thickness of the epidermis covered by basket-weave orthokeratotic keratin. Hematoxylin and Eosin 40x. **(C,D)** Skin sections of the same dogs as in A and B at higher magnification (100x).

doi:10.1371/journal.pgen.1006651.g002

infundibula in all biopsies. The keratin layers were multifocally exfoliating as large scales. The underlying epidermis was mildly hyperplastic. In the biopsy from the inguinal region, the infundibula of the hair follicles were moderately dilated. The histological findings were consistent with a cornification disorder and an inherited non-epidermolytic ichthyosis as possible cause (Fig 2).

Genetic analysis

We sequenced the genome of the affected dog at 31x coverage and called SNVs and small indel variants with respect to the reference genome (CanFam 3.1). We then compared these variants to whole genome sequence data of 288 control dogs of various breeds including 13 German Shepherds not closely related to the affected dog (Table 1). As we did not find any protein-changing variants in the 14 known ichthyosis-associated genes, we hypothesized that the affected dog represented an isolated case of a novel form of ichthyosis. Consequently, we considered both a recessive and a dominant mode of inheritance for the hypothetical mutant allele.

As purebred dogs are maintained in closed populations with a small effective population size and a considerable degree of inbreeding, recessive genetic defects within a breed typically can be traced back to single founders and are mostly found in homozygous state in affected

Table 1. Variants detected by whole genome re-sequencing of the affected dog.

Filtering step[a]	Number of variants
Homozygous variants in the whole genome	3,108,583
Private homozygous variants (absent from 288 control genomes)	797
Protein-changing private homozygous variants (absent from 288 control genomes)	4
Heterozygous variants in the whole genome	2,767,699
Private heterozygous variants (absent from 288 control genomes)	3,202
Protein-changing private heterozygous variants (absent from 288 control genomes)	19

[a] The sequences were compared to the reference genome (CanFam 3.1) from a Boxer.

doi:10.1371/journal.pgen.1006651.t001

dogs. In a first analysis, assuming a recessive mode of inheritance, we therefore searched for homozygous private protein-changing variants in the affected dog. Our automated pipeline detected 4 such homozygous protein-changing variants. However, upon individual visual inspection all 4 variants turned out to be sequencing artifacts. They were either located close to gaps in the reference assembly, in highly repetitive sequences, or in regions with low read coverage (S1 Table).

Given that the described dog was the only case in a litter of seven and ichthyosis had never before been reported in the German Shepherd breed, we hypothesized that a dominant mode of inheritance due to a *de novo* mutation event was more likely than a recessive mode of inheritance. In a second analysis, we therefore filtered for heterozygous private protein-changing variants. Our automated pipeline identified 19 such variants in 13 genes. None of the identified variants was located in a known ichthyosis gene. In order to identify potential *de novo* variants, we obtained whole genome sequences from both parents of the affected dogs. Inspection of the sequencing data for each of the 19 heterozygous candidate variants revealed that only one of them was indeed a *de novo* variant (S1 Table).

This *de novo* variant was a missense variant, c.1052T>C, located in the *ASPRV1* gene encoding "aspartic peptidase, retroviral-like 1" also known as skin aspartic protease (SASPase), which is involved in profilaggrin-to-filaggrin processing [17,18,22]. We performed Sanger sequencing in the affected dog and both parents and confirmed that the variant was absent in both parents (Fig 3A). We experimentally confirmed the correct parentage by an analysis of microsatellite and SNV genotypes in the trio.

The ASPRV1:c.1052T>C variant is predicted to result in the amino acid substitution p.(Leu351Pro). The leucine at position 351 is strictly conserved among different species of placental mammals and only one residue away from one of the major cleavage sites required for auto-activation of the protein (Fig 3B) [17].

Functional confirmation

To assess the putative impact of the ASPRV1 missense variant we performed immunofluorescence staining with anti-ASPRV1 antibodies on skin sections of the affected and a control dog. The ASPRV1 signal in the affected dog showed the expected localization, mainly in the stratum granulosum, but was stronger than in the control dog (Fig 4). As this experiment could not assess whether the detected ASPRV1 protein is functional, we also investigated filaggrin processing by immunofluorescence staining with anti-filaggrin antibodies. This experiment demonstrated an abnormal filaggrin expression pattern in the affected dog (Fig 4). In the affected dog, diffuse staining across epidermal layers (from stratum basale through stratum spinosum to stratum granulosum) and some nuclear staining indicated defective processing of profilaggrin.

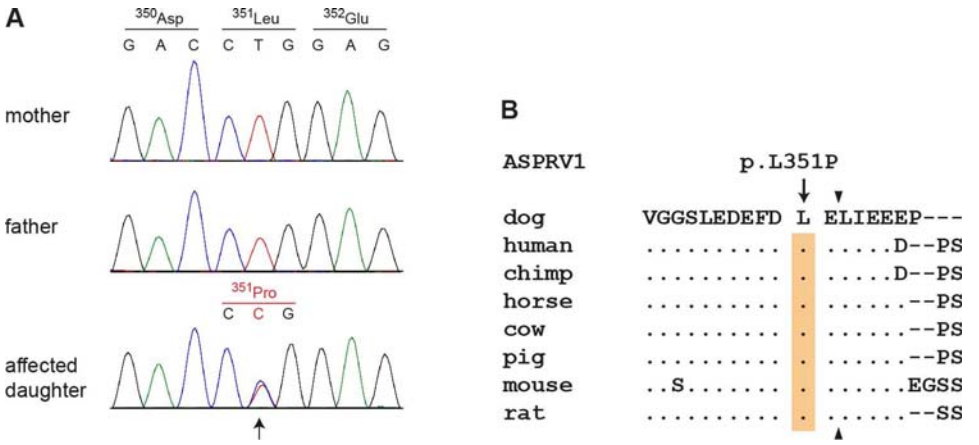


Fig 3. Sanger electropherograms of the ASPRV1:c.1052T>C variant and evolutionary conservation of leucine 351 in the ASPRV1 protein. (A) A genomic ASPRV1 fragment was amplified by PCR and sequenced with the Sanger method. The figure shows genotypes of the affected daughter and both her non-affected parents. The position of the variant is indicated by an arrow. Note that the variant C-allele is only present in the daughter, but absent from both parents indicating a *de novo* mutation event. **(B)** The leucine residue at position 351 of the canine ASPRV1 protein is strictly conserved in several species and located close to the C-terminal auto cleavage site, which is indicated by arrowheads. The multiple alignment was done using accessions XP_013972931.1 (*Canis lupus familiaris*), NP_690005.2 (*Homo sapiens*), XP_525777.1 (*Pan troglodytes*), XP_014586589.1 (*Equus caballus*), XP_003586694.1 (*Bos taurus*), XP_003354829.2 (*Sus scrofa*), NP_080690.2 (*Mus musculus*) and XP_008761336.1 (*Rattus norwegicus*). A full-length alignment of the proteins from selected species is given in S1 Fig.

doi:10.1371/journal.pgen.1006651.g003

Discussion

In the present study we identified a *de novo* missense variant in the canine *ASPRV1* gene in a dog with a novel form of ichthyosis. We provide five arguments supporting the causality of the ASPRV1:c1052T>C variant for the observed ichthyosis.

First, the c.1052T>C variant leads to a non-conservative amino acid exchange p.(Leu351Pro) close to the functionally important auto-cleavage site of the ASPRV1 protease. It is thus conceivable that this specific genetic variant might affect the protein function.

Second, the c.1052T>C was absent from 288 non-affected dogs of different breeds and thus perfectly associated with the disease status.

Third, the affected dog was heterozygous for the variant, but the mutant allele was absent in blood leukocytes of both parents. We therefore confirmed that the variant had arisen by a *de novo* mutation event that must have occurred in either one of the parental germlines or during early embryonic development of the affected dog. While exact numbers on the frequencies of *de novo* mutation events in dogs were not available at the time of this study, an analysis of 10 human trios reported 73 *de novo* mutation events on average per trio [19]. In another study on human *de novo* mutation events, it was shown that only ~1.3% of the *de novo* variants actually represented protein-changing variants [20]. Similar numbers of *de novo* mutation events were observed in cattle [21]. If we assume that these numbers are similar in dogs, one might expect roughly one *de novo* protein-changing variant in any dog on average, which exactly matches our data with one identified protein-changing variant in the affected dog. We deem it unlikely that such an event would coincidentally affect a gene with a known role in filaggrin processing without being causative for the ichthyosis phenotype.

Fourth, we observed a difference in the ASPRV1 protein expression between the affected dog and a control dog. Somewhat surprisingly, the ASPRV1 protein expression was upregulated in the affected dog. Such an upregulation might have been caused by a compensatory

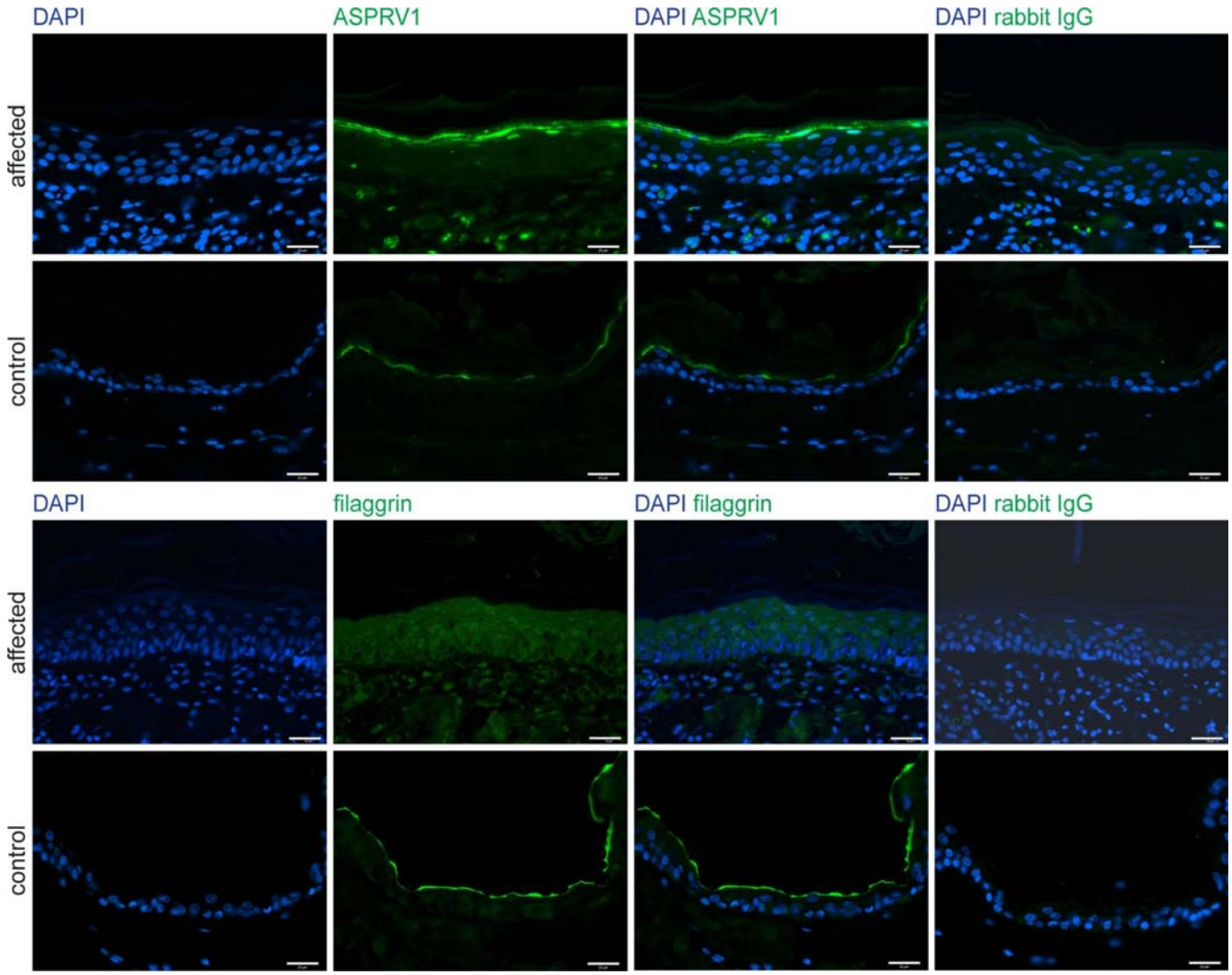


Fig 4. ASPRV1 and filaggrin protein expression. Skin sections of the case and a non-affected control dog were stained for immunofluorescence with an anti-ASPRV1 or anti-filaggrin antibody, respectively. The ASPRV1 signal was stronger in the affected dog than in the control dog. The typical specific expression pattern of filaggrin, a line at the border between stratum granulosum and stratum corneum, was observed in the control dog (bottom panels) but not in the affected dog (third row of panels). Specificity of the antibodies was demonstrated by using rabbit IgG as primary antibody (last column). Scale bars: 25 μm.

doi:10.1371/journal.pgen.1006651.g004

mechanism, if the expressed ASPRV1 protein is non-functional as has been reported for other missense variants [22]. Thus, the increased ASPRV1 quantity is at least compatible with a causal role of the ASPRV1 missense variant in the observed ichthyosis.

Finally and as fifth argument, we experimentally confirmed that the filaggrin protein expression pattern was altered in the affected dog suggesting a defect in profilaggrin-to-filaggrin processing. In our opinion and taken together, these arguments prove the causality of the ASPRV1:c1052T>C variant for the observed ichthyosis beyond any reasonable doubt.

The *ASPRV1* gene and its encoded protease were initially identified in humans and shortly afterwards in mice. ASPRV1 protein expression was only detected in stratified epithelia and was restricted to the stratum granulosum [17,18]. Further studies suggested that ASPRV1 cleaves the linker sequence in profilaggrin. A deficiency of ASPRV1 resulted in a lower level of

stratum corneum hydration [23]. Furthermore, high ASPRV1 protein levels were present in several non-neoplastic skin disorders [17,24].

In transgenic mice, aberrant Asprv1 expression resulted in delayed wound closure [25]. *Asprv1* deficient mice (*Asprv1*$^{-/-}$) in a C57BL/6J background showed characteristic parallel skin wrinkles or lined grooves parallel to the body axis, but were reported to have normal skin histology and did not show any signs of ichthyosis [18]. The skin of *Asprv1*$^{-/-}$ mice in a hairless background (Hos:HR-1) displayed more fine wrinkles and was drier and rougher than in *Asprv1*$^{+/-}$ or *Asprv1*$^{+/+}$ mice [23]. In addition to an increased number of epidermal cell layers and a lower level of stratum corneum hydration, a decreased amount of filaggrin monomers together with an accumulation of aberrantly processed profilaggrin in the lower stratum corneum was observed in these mice. The total amount and the composition of free amino acids was however not significantly different from control mice [23].

Our results clearly showed an altered filaggrin expression pattern in the skin of an affected dog with the *ASPRV1* variant. Thus, similar to Asprv1 deficient mice, profilaggrin processing appeared to be defective in the affected dog. In contrast to the findings by Matsui et al. [23], we did however not detect an accumulation of incompletely processed filaggrin in the stratum corneum, but rather in the stratum spinosum and stratum granulosum. It remains unclear why the *Asprv1*$^{-/-}$ mice did not show an ichthyosis phenotype as the *ASPRV1* mutant dog. Potential explanations include a gain of function effect of the specific canine missense variant or general differences in the homeostasis of the epidermis between mice and dogs.

According to our knowledge, the *ASPRV1* gene has not been associated with any form of ichthyosis in humans. Missense variants in the human *ASPRV1* gene were reported in 5 of 196 Japanese patients with atopic eczema and 2 of 28 control subjects [23]. Two of the identified variants, p.V243A (identified in a control subject) and p.V187I (identified in 3 atopic eczema patients) led to absence or reduction of ASPRV1 activity *in vitro* [23]. Another study failed to find significant associations between *ASPRV1* genetic variants and atopic eczema or clinically dry skin in different cohorts of Caucasian ancestry [26].

In conclusion, with the identification of a dominant *de novo* missense variant in the *ASPRV1* gene of an ichthyotic dog, we present a new candidate gene for ichthyosis. It seems possible that *ASPRV1* variants might also contribute to unsolved human ichthyoses.

Materials and methods

Ethics statement

All animal experiments were performed according to the local regulations. The dogs in this study were examined with the consent of their owners. The study was approved by the "Cantonal Committee For Animal Experiments" (Canton of Bern; permits 22/07, 23/10, and 75/16).

Clinical and histopathological examination

The affected dog was examined by a board certified veterinary dermatologist in a private specialist clinic and followed up after initiating treatment with a rehydrating, anti-seborrheic spray (Ermidra, Ufamed AG, Sursee, Switzerland) and shampoo (Sebomild P, Virbac AG, Glattbrugg, Switzerland).The absence of a similar phenotype in littermates, parents and ancestors was reported by the owner. Skin biopsies (6 mm) of the case were taken from the flank, thigh, shoulder, and inguinal region and fixed in 10% buffered formalin for 24 hours. Biopsies were processed, embedded in paraffin and sectioned at 4 μm. Skin sections were stained with hematoxylin and eosin. The histopathology was performed by board certified pathologists.

DNA isolation and parentage confirmation

We isolated genomic DNA from EDTA blood samples of the affected dog and its parents. We confirmed the parentage by two different approaches: A multiplex PCR with 7 fluorescently labeled microsatellites primer pairs was performed for both parents and the case. Allele sizes were determined on an ABI 3730 capillary sequencer (Life Technologies) and analyzed using the GeneMapper software (Life Technologies). Mendelian transmission of the alleles was confirmed at all 7 loci. The three dogs were additionally genotyped for 173,662 SNVs on the illumina canine_HD chip by GeneSeek/Neogen. We evaluated the parentage using the --genome and--mendel commands in plink 1.07 [27]. Both analyses were in agreement with the assumed parentage (<0.001% genome regions with IBD = 0 for mother and father; 26 Mendel errors for the trio).

Whole genome sequencing of the affected German Shepherd and its parents

Illumina PCR-free TruSeq fragment libraries with insert sizes of 350 bp—400 bp were prepared. For the affected dog, 276 million 2 x 150 bp paired-end reads or 31x coverage were obtained on a HiSeq3000 instrument. The parents were sequenced at 21x coverage. The reads were mapped to the dog reference genome assembly CanFam3.1 and aligned using Burrows-Wheeler Aligner (BWA) version 0.7.5a [28] with default settings. The generated SAM file was converted to a BAM file and the reads were sorted using samtools [29]. Picard tools (http://sourceforge.net/projects/picard/) was used to mark PCR duplicates. To perform local realignments and to produce a cleaned BAM file, we used the Genome Analysis Tool Kit (GATK version 2.4.9, 50) [30]. GATK was also used for base quality recalibration with canine dbsnp data as training set. The sequence data were deposited under the study accession PRJEB16012 at the European Nucleotide Archive. The sample accessions are SAMEA4506895 for the case (DS043), SAMEA72802918 for the sire (DS053) and SAMEA72802168 for the dam (DS051).

Variant calling

Putative SNVs were identified in each sample individually using GATK HaplotypeCaller in gVCF mode, and subsequently genotyped per-chromosome and genotyped across all samples simultaneously [31]. Filtering was performed using the variant filtration module of GATK. To predict the functional effects of the called variants, SnpEFF [32] software together with the ENSEMBL (version 80) annotation CanFam 3.1 was used. We additionally visually inspected the short read alignments of the functional candidate genes *ABCA12*, *ALOX12B*, *ALOXE3*, *CERS3*, *CYP4F22*, *FLG*, *KRT1*, *KRT10*, *KRT2*, *LIPN*, *NIPAL4*, *PNPLA1*, *STS*, and *TGM1* in the integrative genome viewer [33] to exclude any structural variants in these genes. We also inspected the *CLDN1* gene in the same manner. For variant filtering we used 288 control genomes, which were either publicly available [34] or produced during other projects of our group. A detailed list of these control genomes is given in S2 Table.

Gene analysis

We used the dog CanFam 3.1 reference genome assembly for all analyses. Numbering within the canine *ASPRV1* gene corresponds to the accessions XM_014117456.1 (mRNA) and XP_013972931.1 (protein). Numbering within the human *ASPRV1* gene corresponds to the accessions NM_152792.2 (mRNA) and NP_690005.2 (protein).

Sanger sequencing

We used Sanger sequencing to confirm the candidate variant c.1052T>C in *ASPRV1* in the affected dog and its absence in both parents. A 370 bp fragment containing the variant was PCR amplified from genomic DNA using AmpliTaq Gold 360 Master Mix (Life Technologies) and the primers ACCCCAGGGACAGATTAAGG and AGCTGAAGCTGAAGGCAGAG. After treatment with shrimp alkaline phosphatase and endonuclease I, PCR products were directly sequenced on an ABI 3730 capillary sequencer (Life Technologies). We analyzed the Sanger sequence data using the software Sequencher 5.1 (GeneCodes).

Immunofluorescence staining and fluorescence microscopy

Immunofluorescence staining was performed on formalin-fixed paraffin-embedded skin sections of an age- and sex-matched control dog and the affected dog with some adaptations as described previously [35]. Briefly, tissue sections were deparaffinized using xylene. For ASPRV1 staining, antigens were retrieved in a microwave oven for 20 min at 95˚C in sodium citrate buffer (10 mM sodium citrate, pH 6.0). Blocking was performed for 1.5 hours at room temperature (10% goat serum, 1% BSA, 0.1% Triton X-100; 5% cold fish gelatin in PBS). Tissue sections were incubated with a polyclonal rabbit anti-ASPRV1 antibody (1:250, NBP2-33981, Novus Biological) overnight at 4˚C and with the secondary goat anti-rabbit Alexa Fluor 488 nm antibody (1:1000, Abcam) for 1 hour at room temperature. DNA was stained with 4',6-diamidino-2-phenylindole (DAPI) contained in Vectashield Antifade Mounting Medium (Vector Laboratories). For filaggrin staining, antigens were retrieved in a pressure cooker for 15 min in Tris buffer (100 mM) with 5% urea. Blocking was performed for 1.5 hours at room temperature (10% goat serum, 1% BSA, 0.1% Triton X-100; 5% cold fish gelatin in PBS). Tissue sections were incubated with a polyclonal rabbit anti-filaggrin antibody (1:250, PRB-417P-100, Covance) overnight at 4˚C and with the secondary goat anti-rabbit Alexa Fluor 488 nm antibody (1:1000, Abcam) for 1 hour at room temperature. DNA was stained with DAPI (1:1000, Sigma Aldrich). Tissue sections serving as negative controls were incubated with rabbit IgG serum. Images were taken with a Nikon Eclipse 80i fluorescence microscope using a Plan Flour x40/10 oil-immersion objective and excitation wavelength of 393 and 488nm. Pictures were captured and further processed using Improvision Open Lab 5.5.2. software.

Supporting information

S1 Fig. ASPRV1 protein alignment.
(PDF)

S1 Table. Private protein-changing variants in the affected dog.
(XLSX)

S2 Table. Control dogs used for whole genome sequencing.
(XLSX)

Acknowledgments

The authors are grateful to the dog owners who donated samples and shared pedigree data of their dogs. We thank Nathalie Besuchet Schmutz, Muriel Fragnière, Fabiana Jakob, and Sabrina Schenk for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing the high-throughput sequencing experiments, and the Vital-IT high-performance computing center of the Swiss Institute of Bioinformatics for performing computationally intensive tasks (http://www.vital-it.ch/). We thank the Dog Biomedical Variant Database

Consortium (Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Cord Drögemüller, Oliver Forman, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Vidhya Jagannathan, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Anita Oberbauer, Jeffrey Schoenebeck, Sheila Schmutz, Claire Wade) for sharing whole genome sequencing data from control dogs. We also acknowledge all canine researchers who deposited dog whole genome sequencing data into public databases.

Author Contributions

Conceptualization: EJM PR MMW TL.

Data curation: VJ.

Funding acquisition: EJM PR MMW TL.

Investigation: AB DPW AG KT VJ BSS DJW.

Methodology: AB DPW AG VJ BSS.

Resources: KT ED.

Supervision: AG EJM PR MMW TL.

Writing – original draft: AB KT DJW TL.

Writing – review & editing: AB DPW AG KT VJ BSS DJW EJM PR MMW TL.

References

1. Takeichi T, Akiyama M (2016) Inherited ichthyosis: Non-syndromic forms. J Dermatol 43: 242–251. doi: 10.1111/1346-8138.13243 PMID: 26945532
2. Oji V, Tadini G, Akiyama M, Blanchet Bardon C, Bodemer C, Bourrat E, et al. (2010) Revised nomenclature and classification of inherited ichthyoses: results of the First Ichthyosis Consensus Conference in Soreze 2009. J Am Acad Dermat 63: 607–641.
3. Wells RS, Kerr CB (1966) Clinical features of autosomal dominant and sex-linked ichthyosis in an English population. Br Med J 1: 947–950. PMID: 20790920
4. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, et al. (2006) Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nat Genet 38: 337–342. doi: 10.1038/ng1743 PMID: 16444271
5. Candi E, Schmidt R, Melino G (2005) The cornified envelope: a model of cell death in the skin. Nat Rev Mol Cell Biol 6: 328–340. doi: 10.1038/nrm1619 PMID: 15803139
6. Dale BA, Resing KA, Lonsdale-Eccles JD (1985) Filaggrin: a keratin filament associated protein. Ann New York Acad Sci 455: 330–342.
7. Presland RB, Haydock PV, Fleckman P, Nirunsuksiri W, Dale BA (1992) Characterization of the human epidermal profilaggrin gene. Genomic organization and identification of an S-100-like calcium binding domain at the amino terminus. J Biol Chem 267: 23772–23781. PMID: 1429717
8. Hoste E, Kemperman P, Devos M, Denecker G, Kezic S, Yau N, et al. (2011) Caspase-14 is required for filaggrin degradation to natural moisturizing factors in the skin. J Invest Dermatol 131: 2233–2241. doi: 10.1038/jid.2011.153 PMID: 21654840
9. Basler E, Grompe M, Parenti G, Yates J, Ballabio A (1992) Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis. Am J Hum Genet 50: 483–491. PMID: 1539590
10. Lemke JR, Kernland-Lang K, Hörtnagel K, Itin P (2014) Monogenic human skin disorders. Dermatol 229:55–64.
11. Grall A, Guaguere E, Planchais S, Grond S, Bourrat E, Hausser I, et al. (2012) PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nat Genet 44: 140–147. doi: 10.1038/ng.1056 PMID: 22246504
12. Tamamoto-Mochizuki C, Banovic F, Bizikova P, Laprais A, Linder KE, Olivry T (2016) Autosomal recessive congenital ichthyosis due to PNPLA1 mutation in a golden retriever-poodle cross-bred dog and the effect of topical therapy. Vet Dermatol 27: 306–e75. doi: 10.1111/vde.12323 PMID: 27237723

13. Credille KM, Barnhart KF, Minor JS, Dunstan RW (2005) Mild recessive epidermolytic hyperkeratosis associated with a novel keratin 10 donor splice-site mutation in a family of Norfolk terrier dogs. Br J Dermatol 153: 51–58. doi: 10.1111/j.1365-2133.2005.06735.x PMID: 16029326

14. Mauldin EA, Wang P, Evans E, Cantner CA, Ferracone JD, Credille KM et al. (2015) Autosomal recessive congenital ichthyosis in American Bulldogs is associated with NIPAL4 (ICHTHYIN) deficiency. Vet Pathol 52: 654–662. doi: 10.1177/0300985814551425 PMID: 25322746

15. Credille KM, Minor JS, Barnhart KF, Lee E, Cox ML, Tucker KA, et al. (2009) Transglutaminase 1-deficient recessive lamellar ichthyosis associated with a LINE-1 insertion in Jack Russell terrier dogs. Br J Dermatol 161: 265–272. doi: 10.1111/j.1365-2133.2009.09161.x PMID: 19438474

16. Mauldin EA (2013) Canine ichthyosis and related disorders of cornification. Vet Clin North Am Small Anim Pract 43: 89–97. doi: 10.1016/j.cvsm.2012.09.005 PMID: 23182326

17. Bernard D, Mehul B, Thomas-Collignon A, Delattre C, Donovan M, Schmidt R (2005) Identification and characterization of a novel retroviral-like aspartic protease specifically expressed in human epidermis. J Invest Dermatol 125: 278–287. doi: 10.1111/j.0022-202X.2005.23816.x PMID: 16098038

18. Matsui T, Kinoshita-Ida Y, Hayashi-Kisumi F, Hata M, Matsubara K, Chiba M, et al. (2006) Mouse homologue of skin-specific retroviral-like aspartic protease involved in wrinkle formation. J Biol Chem 281: 27512–27525. doi: 10.1074/jbc.M603559200 PMID: 16837463

19. Besenbacher S, Liu S, Izarzugaza JM, Grove J, Belling K, Bork-Jensen J, et al. (2015) Novel variation and *de novo* mutation rates in population-wide *de novo* assembled Danish trios. Nat Comm 6: 5969.

20. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, et al. (2012) Rate of *de novo* mutations and the importance of father's age to disease risk. Nature 488: 471–475. doi: 10.1038/nature11396 PMID: 22914163

21. Harland C, Charlier C, Karim L, Cambisano N, Deckers M, Mullaart E, et al. (2016) Frequency of mosaicism points towards mutation-prone early cleavage cell divisions. bioRxiv.

22. Matsuda A, Kosugi S (1997) A homozygous missense mutation of the sodium/iodide symporter gene causing iodide transport defect. J Clin Endocrinol Metab 82: 3966–3971. doi: 10.1210/jcem.82.12.4425 PMID: 9398697

23. Matsui T, Miyamoto K, Kubo A, Kawasaki H, Ebihara T, Hata K, et al. (2011) SASPase regulates stratum corneum hydration through profilaggrin-to-filaggrin processing. EMBO Mol Med 3: 320–333. doi: 10.1002/emmm.201100140 PMID: 21542132

24. Rhiemeier V, Breitenbach U, Richter KH, Gebhardt C, Vogt I, Hartenstein B, et al. (2006) A novel aspartic proteinase-like gene expressed in stratified epithelia and squamous cell carcinoma of the skin. Am J Pathol 168: 1354–1364. doi: 10.2353/ajpath.2006.050871 PMID: 16565508

25. Hildenbrand M, Rhiemeier V, Hartenstein B, Lahrmann B, Grabe N, Angel P, et al. (2010) Impaired skin regeneration and remodeling after cutaneous injury and chemically induced hyperplasia in taps-transgenic mice. J Invest Dermatol 130: 1922–1930. doi: 10.1038/jid.2010.54 PMID: 20237492

26. Sandilands A, Brown SJ, Goh CS, Pohler E, Wilson NJ, Campbell LE, et al. (2012) Mutations in the SASPase gene (ASPRV1) are not associated with atopic eczema or clinically dry skin. J Invest Dermatol 132: 1507–1510. doi: 10.1038/jid.2011.479 PMID: 22318384

27. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575. doi: 10.1086/519795 PMID: 17701901

28. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760. doi: 10.1093/bioinformatics/btp324 PMID: 19451168

29. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987–2993. doi: 10.1093/bioinformatics/btr509 PMID: 21903627

30. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297–1303. doi: 10.1101/gr.107524.110 PMID: 20644199

31. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 43:11.0.1–33.

32. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6: 80–92. doi: 10.4161/fly.19695 PMID: 22728672

33. Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14: 178–192. doi: 10.1093/bib/bbs017 PMID: 22517427

34. Bai B, Zhao WM, Tang BX, Wang YQ, Wang L, Zhang Z, et al. (2015) DoGSD: the dog and wolf genome SNP database. Nucleic Acids Res 43 (Database issue): D777–783. doi: 10.1093/nar/gku1174 PMID: 25404132

35. Chervet L, Galichet A, McLean WH, Chen H, Suter MM, Roosje PJ, et al. (2010) Missing C-terminal filaggrin expression, NFkappaB activation and hyperproliferation identify the dog as a putative model to study epidermal dysfunction in atopic dermatitis. Exp Dermatol. 19: e343–346. doi: 10.1111/j.1600-0625.2010.01109.x PMID: 20626465

MKLN1 splicing defect in dogs with lethal acrodermatitis

Journal:	PLoS Genetics
Manuscript status:	published
Contributions:	genetic analyses, original draft, review and editing of manuscript

PLOS GENETICS

RESEARCH ARTICLE

MKLN1 splicing defect in dogs with lethal acrodermatitis

Anina Bauer[1,2], Vidhya Jagannathan[1,2], Sandra Högler[3], Barbara Richter[3], Neil A. McEwan[4], Anne Thomas[5], Edouard Cadieu[6], Catherine André[6], Marjo K. Hytönen[7,8,9], Hannes Lohi[7,8,9], Monika M. Welle[2,10], Petra Roosje[2,11], Cathryn Mellersh[12], Margret L. Casal[13], Tosso Leeb[1,2]*

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **2** DermFocus, University of Bern, Bern, Switzerland, **3** Department of Pathobiology, Institute of Pathology and Forensic Veterinary Medicine, University of Veterinary Medicine Vienna, Vienna, Austria, **4** Department of Small Animal Clinical Sciences, The University of Liverpool, Leahurst Campus, Neston, Cheshire, United Kingdom, **5** Antagene, Animal Genetics Laboratory, La Tour de Salvagny, France, **6** Institut de Génétique et Développement de Rennes (IGDR), CNRS-UMR6290, Université Rennes1, Rennes, France, **7** Department of Veterinary Biosciences, University of Helsinki, Helsinki, Finland, **8** Research Programs Unit, Molecular Neurology, University of Helsinki, Helsinki, Finland, **9** Folkhälsan Institute of Genetics, University of Helsinki, Helsinki, Finland, **10** Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **11** Division of Clinical Dermatology, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern,Bern, Switzerland, **12** Kennel Club Genetics Centre, Animal Health Trust, Kentford, Newmarket, Suffolk, United Kingdom, **13** Section of Medical Genetics, Department of Clinical Sciences & Advanced Medicine, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

* Tosso.Leeb@vetsuisse.unibe.ch

OPEN ACCESS

Citation: Bauer A, Jagannathan V, Högler S, Richter B, McEwan NA, Thomas A, et al. (2018) *MKLN1* splicing defect in dogs with lethal acrodermatitis. PLoS Genet 14(3): e1007264. https://doi.org/10.1371/journal.pgen.1007264

Editor: Judith Fischer, University Medical Center Freiburg, GERMANY

Received: December 20, 2017

Accepted: February 21, 2018

Published: March 22, 2018

Data Availability Statement: The Illumina canine HD array SNV genotypes are available at https://www.animalgenome.org/repository/pub/BERN2017.1208/. The whole genome sequences are available at the European Nuclotide Archive. S2 Table lists accessions of all 192 genomes used for the study. All other relevant data are within the paper and its Supporting Information files.

Funding: Parts of this work were funded by grants from the Swiss National Science Foundation (CRSII3_160738/1; www.snf.ch) and Albert-Heim Foundation (project no. 105; http://www.albert-

Abstract

Lethal acrodermatitis (LAD) is a genodermatosis with monogenic autosomal recessive inheritance in Bull Terriers and Miniature Bull Terriers. The LAD phenotype is characterized by poor growth, immune deficiency, and skin lesions, especially at the paws. Utilizing a combination of genome wide association study and haplotype analysis, we mapped the LAD locus to a critical interval of ~1.11 Mb on chromosome 14. Whole genome sequencing of an LAD affected dog revealed a splice region variant in the *MKLN1* gene that was not present in 191 control genomes (chr14:5,731,405T>G or *MKLN1*:c.400+3A>C). This variant showed perfect association in a larger combined Bull Terrier/Miniature Bull Terrier cohort of 46 cases and 294 controls. The variant was absent from 462 genetically diverse control dogs of 62 other dog breeds. RT-PCR analysis of skin RNA from an affected and a control dog demonstrated skipping of exon 4 in the *MKLN1* transcripts of the LAD affected dog, which leads to a shift in the *MKLN1* reading frame. *MKLN1* encodes the widely expressed intracellular protein muskelin 1, for which diverse functions in cell adhesion, morphology, spreading, and intracellular transport processes are discussed. While the pathogenesis of LAD remains unclear, our data facilitate genetic testing of Bull Terriers and Miniature Bull Terriers to prevent the unintentional production of LAD affected dogs. This study may provide a starting point to further clarify the elusive physiological role of muskelin 1 *in vivo*.

heim-stiftung.ch) to TL; starting grant 260997 from the European Research Council (https://erc.europa.eu/) and a grant from the Jane and Aatos Erkko Foundation (http://jaes.fi/en/) to HL; grant ANR-11-INBS-0003 from the Agence de la Recherche (http://www.agence-nationale-recherche.fr) to CA; grant OD 010939 from the National Institutes of Health (https://www.nih.gov/) and ACORN grant #01746A from the American Kennel Club Canine Health Foundation (http://www.akcchf.org/) to MLC. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: AT is affiliated with a diagnostic lab marketing genetic tests for dogs. HL owns shares of a diagnostic testing lab marketing genetic tests for dogs. CM is affiliated with a diagnostic lab marketing genetic tests for dogs. MLC is affiliated with a diagnostic lab marketing genetic tests for dogs. TL is an associate editor of PLoS Genetics.

Author summary

Lethal acrodermatitis (LAD) is an autosomal recessive hereditary disease in dogs. It is characterized by poor growth, immune deficiency and characteristic skin lesions of the paws and of the face. We mapped the LAD locus to a ~1.11 Mb segment on canine chromosome 14. Whole genome sequence data of an LAD affected dog and 191 controls revealed a candidate causative variant in the *MKLN1* gene, encoding muskelin 1. The identified variant, a single nucleotide substitution, *MKLN1*:c.400+3A>C, altered the 5'-splice site at the beginning of intron 4. We experimentally confirmed that this variant leads to complete skipping of exon 4 in the *MKLN1* mRNA in skin. Various cellular functions have been postulated for muskelin 1 including roles in intracellular transport processes, cell morphology, cell spreading, and cell adhesion. Our data from dogs reveal a novel *in vivo* role for muskelin 1 that is related to the immune system and skin. *MKLN1* thus represents a novel candidate gene for human patients with unsolved acrodermatitis and/or immune deficiency phenotypes. LAD affected dogs may serve as models to gain more insights into the function of muskelin 1.

Introduction

Acrodermatitis enteropathica in humans (OMIM #201100) is an inherited disorder of zinc metabolism. Affected patients display an inflammatory rash, diarrhea and a general failure to thrive [1–3]. This disease is caused by variants in the *SLC39A4* gene encoding a zinc transporter that mediates the uptake of dietary zinc in the gut. Clinical signs in patients will ameliorate or even resolve upon oral supplementation with zinc [4]. A similar *SLC39A4* associated hereditary zinc deficiency exists in cattle [5].

In Bull Terriers, a related phenotype termed lethal acrodermatitis (LAD) has been reported in the scientific literature as early as 1986 [6]. LAD is inherited as a monogenic autosomal recessive trait. Affected puppies show characteristic skin lesions on the feet and on the face, diarrhea, bronchopneumonia, and a failure to thrive. The skin lesions consist of erythema and tightly adherent scales, erosions or ulcerations with crusts involving primarily the feet, distal limbs, elbows, hocks, and muzzle. Later on, hyperkeratosis of the footpads and deformation of the nails occur. LAD affected dogs also show a coat color dilution in pigmented skin areas. An abnormally arched hard palate impacted with decayed, malodorous food is a characteristic clinical marker for the disease (Fig 1) [6–8].

LAD dogs are immunodeficient with a reduction in serum IgA levels and frequently suffer from skin infections with *Malassezia* or *Candida* [9,10]. LAD manifests clinically in the first weeks of life. Affected puppies typically die before they reach an age of two years, either due to infections such as bronchopneumonia or because they are euthanized when their paw pad lesions become very severe and painful. They grow slower than their non-affected littermates and at the age of one year have about half the body weight and size of an unaffected dog [8]. Some, but not all studies found reduced levels of zinc in the serum of LAD affected dogs [6,8,11]. In contrast to acrodermatitis enteropathica in humans, oral or intravenous supplementation of zinc does not lead to an improvement of the clinical signs in LAD affected dogs [6]. A proteomic analysis reported changes related to inflammatory response in the liver of LAD affected puppies [12].

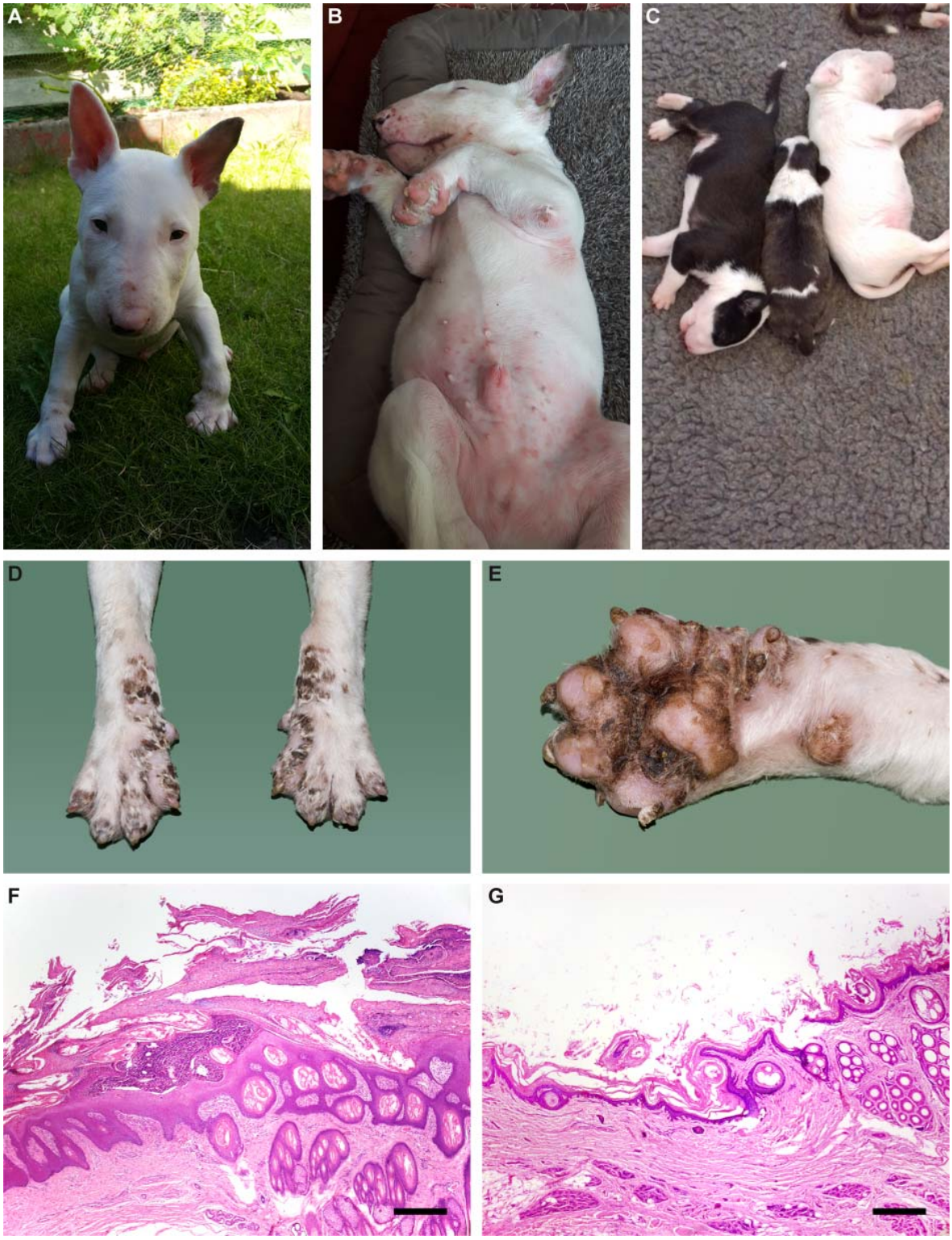


Fig 1. LAD phenotype. **(A)** Inflammatory skin lesions in the face of an affected Bull Terrier. **(B)** Similar lesions in the inguinal region. **(C)** LAD affected puppy in the middle of two non-affected littermates. A pronounced growth delay and a subtle coat color dilution are visible. **(D, E)** Fore paws of an LAD affected Bull Terrier puppy at necropsy. Symmetrical scaling and crusting of the skin including interdigital areas and foot pads is visible **(F, G)** Histopathological micrographs of the junction of interdigital haired skin and digital pad from an affected Bull Terrier

puppy **(F)** and a control dog **(G)**. Marked thickening of the epidermis, excessive layers of non-cornifying epithelium and a large pustule are evident in the affected dog. Hematoxylin-eosin, bar = 400 μm.

https://doi.org/10.1371/journal.pgen.1007264.g001

In the present study, we performed a genome-wide association study (GWAS) followed by a whole genome sequencing approach to unravel the causative genetic variant for LAD in Bull Terriers and Miniature Bull Terriers.

Results

Mapping of the LAD locus

We performed a GWAS with genotypes from 78 Bull Terriers and Miniature Bull Terriers. After quality control, the pruned dataset consisted of 22 LAD cases, 48 controls and 76,419 markers. We obtained a single strong association signal with 57 markers exceeding the Bonferroni-corrected genome-wide significance threshold after adjustment for genomic inflation ($P_{Bonf.} = 6.5 \times 10^{-7}$). All significantly associated markers were located on chromosome 14 within an interval spanning from 0.9 Mb– 10.6 Mb. The three top-associated markers all had a P-value of 1.4×10^{-9} and were located between 5.2 Mb– 5.9 Mb on chromosome 14 (Fig 2).

To narrow down the identified region, we visually inspected the genotypes of the cases to perform autozygosity mapping. We searched for homozygous regions with allele sharing and found one region of ~1.11 Mb, which was shared between all 22 cases. The critical interval for the causative LAD variant corresponded to the interval between the first flanking heterozygous markers on either side or chr14:5,248,244–6,355,383 (CanFam 3.1 assembly).

Identification of a candidate causative variant

We sequenced the genome of an affected Bull Terrier at 24x coverage and called single nucleotide variants (SNVs) and small indel variants with respect to the reference genome (CanFam 3.1). We then compared these variants to whole genome sequence data of 3 wolves and 188 control dogs from genetically diverse breeds. This analysis identified five private homozygous variants in the critical interval in the affected dog (Table 1, S1 Table).

Four of these five variants were intergenic and classified as "modifier" by the SNPeff software. The remaining fifth variant was located within the 5'-splice site of intron 4 of the *MKLN1* gene and its SNPeff impact prediction was "low". The formal designation of this variant is chr14:5,731,405T>G or *MKLN1*:c.400+3A>C (S1 Fig).

We confirmed the presence of this variant by Sanger sequencing (Fig 3). As *MKLN1*:c.400+3A>C represented the only plausible candidate causative variant, we genotyped 251 Bull Terriers, 89 Miniature Bull Terriers, and 462 dogs from 62 other breeds for this variant (Table 2). The variant showed perfect association with the LAD phenotype in Bull Terriers and Miniature Bull Terriers ($P_{Fisher} = 4.8 \times 10^{-58}$). All 46 available cases were homozygous for the variant, whereas the unaffected dogs were either homozygous wildtype or heterozygous. The test dogs included a subset of unaffected Bull Terriers and Miniature Bull Terriers from Finland, which were not specifically collected for this study and therefore considered representative for the general population. The 166 Finnish dogs contained 37 heterozygous dogs (22%). The variant was not found in any of the tested dogs from other breeds.

Functional confirmation

To assess the putative impact of the *MKLN1* variant on splicing, we analyzed the frequency of the wildtype and mutant sequence motifs in a compilation of 186,630 human 5'-splice sites [13,14]. The canine wildtype sequence TAGgtaagg was identical to the sequence of 276 human 5'-splice

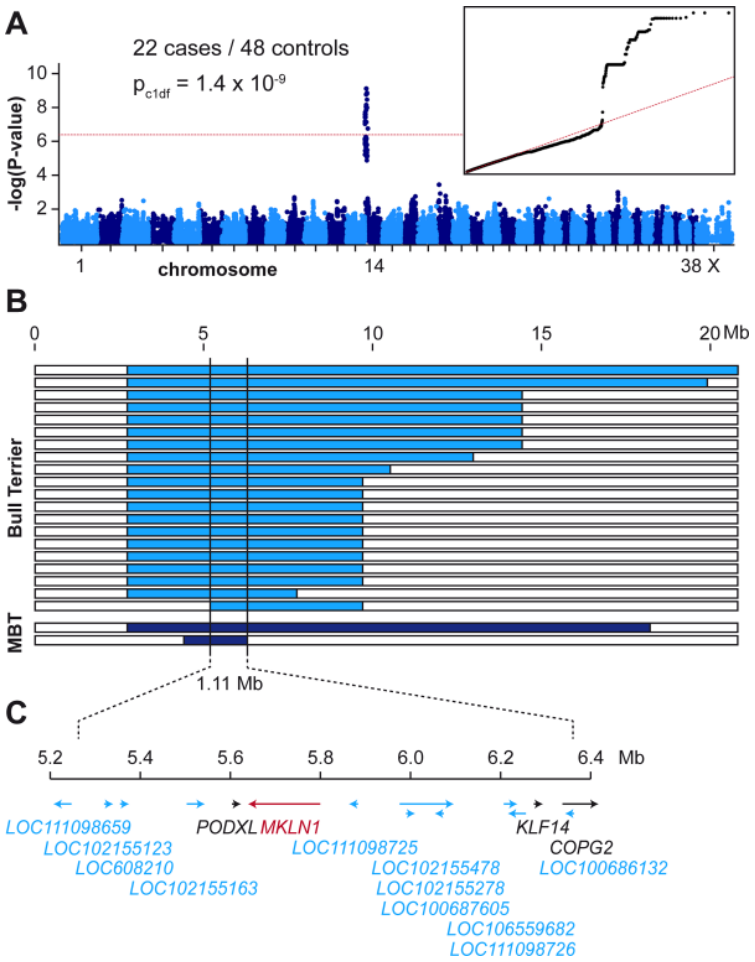


Fig 2. Mapping of the LAD locus. (A) A GWAS was performed in a cohort of 22 LAD cases and 48 controls. The Manhattan plot shows a single significant signal at the beginning of chromosome 14. The red line indicates the Bonferroni significance threshold (P_{Bonf} = 6.5 x 10^{-7}). The quantile-quantile (QQ) plot in the inset shows the observed versus expected–log(p) values. The straight red line in the QQ plot indicates the distribution of p-values under the null hypothesis. The deviation of p-values at the right side indicates that these markers are stronger associated with the trait than it would be expected by chance. **(B)** Haplotype analysis in the 22 LAD cases. Each horizontal bar represents the chromosome 14 haplotypes of one dog. Twenty Bull Terriers and two Miniature Bull Terriers (MBT) had large homozygous intervals with allele sharing on chromosome 14 indicated in blue. The homozygous haplotype segment shared between all 22 dogs spanned ~1.11 Mb. The critical interval for the causative LAD variant corresponded to the interval between the first flanking heterozygous markers on either side or chr14:5,248,244–6,355,383 (CanFam 3.1 assembly). **(C)** Gene annotation for the critical interval. The NCBI annotation release 105 listed 4 protein coding genes (indicated in black or red) and 11 genes for non-coding RNAs (indicated in blue).

https://doi.org/10.1371/journal.pgen.1007264.g002

Table 1. Variants detected by whole genome re-sequencing of an LAD affected dog.

Filtering step[a]	Number of variants
Homozygous variants in the whole genome	3,061,192
Homozygous variants in the 1.11 Mb critical interval on chromosome 14	1,445
Private homozygous variants (absent from 191 control genomes) in critical interval	5

[a] The sequences were compared to the reference genome (CanFam 3.1) from a Boxer.

https://doi.org/10.1371/journal.pgen.1007264.t001

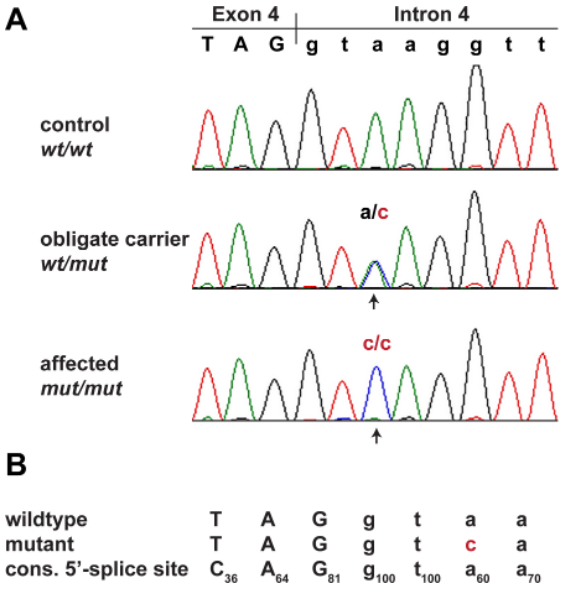


Fig 3. Sanger confirmation of the MKLN1:c.400+3A>C variant. (A) Electropherograms from dogs with the three different genotypes. **(B)** Wildtype and mutant allele compared to the consensus sequence for the human U2 GT-AG type 5'-splice sites [13]. Subscript numbers in the consensus sequence indicate the percentage of the respective conserved nucleotide in 183,682 investigated human 5'-splice site motifs of the U2 GT-AG type. The additional difference to the optimal consensus in the U1 spliceosomal RNA recognition site in the mutant allele is highlighted in red. In human 5'-splice sites the most frequent base at position 3 is an A (60%). G is also common at this position (35%), while C and T are both rare (<3%) [13].

https://doi.org/10.1371/journal.pgen.1007264.g003

sites, while the mutant sequence motif TAGgtcagg occurred in only 3 human 5'-splice sites. The very low frequency of the mutant sequence motif suggested that *MKLN1*:c.400+3A>C might affect the efficacy of the splicing process. Several other pathogenic A>C transversions at 5'-splice sites' position +3 with subsequent exon skipping have been described in the literature [15–17].

We experimentally analyzed *MKLN1* transcripts in skin RNA from an LAD affected dog with the homozygous mutant C/C genotype in comparison to a healthy control dog (A/A genotype). RT-PCR with primers located at the exon 2/3 and exon 5/6 boundaries yielded a cDNA fragment of the expected size in the control dog, but not in the LAD affected dog. In the LAD affected dog, a very clean cDNA amplicon lacking exon 4 was obtained. This experiment demonstrated a complete skipping of exon 4 in *MKLN1* transcripts as consequence of the genomic *MKLN1*:c.400+3A>C variant (r.312_400del89; Fig 4). If translated, the mutant transcript was predicted to result in a severely truncated protein containing only the first 105 of a total of 735 amino acids of the wildtype protein (p.(Gly105SerfsTer10); S2 Fig).

Table 2. Association of the MKLN1:c.400+3A>C genotype with the LAD phenotype.

MKLN1:c.400+3A>C genotype	A/A	A/C	C/C
Cases, Bull Terrier (n = 41)	-	-	41
Cases, Miniature Bull Terrier (n = 5)	-	-	5
Controls, Bull Terrier (n = 210)	132	78	-
Controls, Miniature Bull Terrier (n = 84)	59	25	-
Controls, other breeds (n = 462)	462	-	-

https://doi.org/10.1371/journal.pgen.1007264.t002

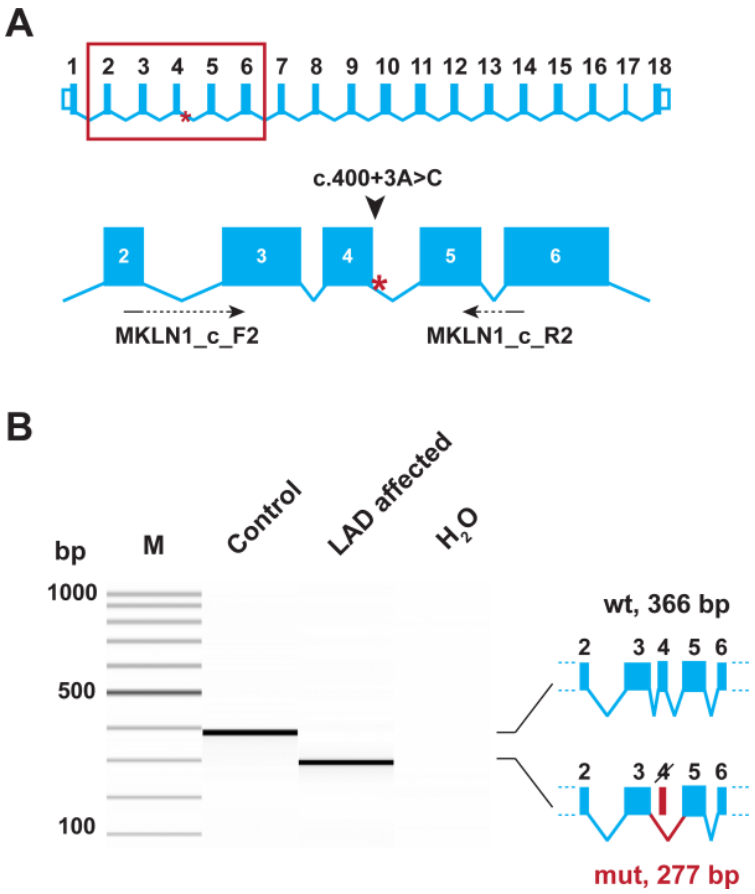


Fig 4. Experimental verification of the MKLN1 splice defect. (A) The genomic organization of the *MKLN1* gene. Exons 2–6 are enlarged and the position of the primers used for RT-PCR is indicated. **(B)** RT-PCR was performed using skin cDNA from a control and an LAD affected Bull Terrier. The picture shows a Fragment Analyzer gel image of the experiment. In the control animal, only the expected 366 bp product is visible. In the LAD affected dog, a 277 bp product representing a transcript lacking exon 4 is visible. The identity of the bands was verified by Sanger sequencing. Thus, the *MKLN1*:c.400+3A>C variant leads to complete skipping of exon 4 (*MKLN1*:r.312_400del89).

https://doi.org/10.1371/journal.pgen.1007264.g004

Discussion

In the present study we identified a splice defect in the canine *MKLN1* gene in Bull Terriers with LAD. The combination of GWAS and haplotype analysis localized the causative variant to a relatively small chromosomal region with only a few characterized genes including *MKLN1*. The splice region variant in *MKLN1* was the only plausible variant within this critical interval that showed the expected genotype concordance with the LAD phenotype in a large cohort of more than 300 Bull Terriers and ~500 dogs from other breeds.

The identified *MKLN1*:c.400+3A>C variant resulted in exon 4 skipping and a frameshift as 89 nucleotides were missing from the mutant transcripts. It therefore seems likely that mutant transcripts are degraded by nonsense-mediated mRNA decay. Considering the strong genetic association of the variant with the phenotype and the fact that we demonstrated a functional defect on the *MKLN1* transcript level, we think that our data strongly suggest the causality of the *MKLN1*:c.400+3A>C variant for LAD in Bull Terriers and Miniature Bull Terriers.

MKLN1 encodes the widely expressed intracellular protein muskelin 1, also known as TWA2. The function of muskelin 1 is only partially understood. It was originally described as

a protein that mediates adhesive and cell-spreading responses to thrombospondin 1, an extracellular matrix adhesion molecule [18]. However, different studies suggested that the function of muskelin 1 goes beyond this pathway and are also supported by the fact that muskelin 1, which has homologs in invertebrates and even fission yeast, evolved earlier than the vertebrate-specific thrombospondin 1 [19,20]. Muskelin 1 is a multidomain protein with an N-terminal discoidin domain, a LisH / CTLH tandem domain, and six C-terminal Kelch repeats, which forms homotetramers [21]. The LisH domain was shown to be crucial for muskelin 1 dimerization and cytoplasmic localization, and, together with the head-to-tail interaction via the discoidin domain, also for the tetramerization of muskelin 1 [20,21].

Consistent with its multidomain structure and ubiquitous expression, diverse binding partners have been reported for muskelin 1. It binds prostaglandin EP3 receptor isoform α [22] and heme-oxidase 1, which counteracts inflammatory and reactive oxygen species induced damage [23]. It is part of the CTLH complex, the homolog of yeast E3 ubiquitin ligase, where it binds to RanBPM and Twa I [24–26] and interacts with the cardiogenic transcription factor TBX-20 [27]. In the rat lens, muskelin 1 is a substrate of Cdk5 and interacts with the Cdk5 activator p39 [28]. Also in lens, it was shown that p39 links muskelin 1 to myosin II and stress fibers [29].

Mkln1$^{-/-}$ knockout mice are viable and do not have skin lesions comparable to those in Bull Terriers with LAD. However, they exhibit a subtle coat color dilution phenotype similar to that seen in LAD affected dogs. In these *Mkln1*$^{-/-}$ knockout mice, muskelin 1 was identified as a protein required for $GABA_A$ receptor endocytosis and trafficking in neurons via direct interaction with the α1 subunit of $GABA_A$ receptors and the motor proteins dynein and myosin VI. The dilute coat color of *Mkln1*$^{-/-}$ knockout mice suggested that muskelin is a trafficking factor involved in several different intracellular transport processes, possibly including melanosome transport [30].

The lacking skin lesions in *Mkln1*$^{-/-}$ knockout mice raise the questions whether muskelin 1 depletion does not result in disease in mice; whether their clinical signs would only manifest at a (much) older age; or whether the sterile environment of the laboratory animals prevented infections and thus the development of skin lesions. In the latter case, LAD would be a primary immunodeficiency disorder, in agreement with the observation of lower IgA levels and higher susceptibility to microbial infection in LAD affected dogs [8–10, 12]. Given the diverse known protein-protein interactions of muskelin 1, it is however likely that absence of muskelin 1 leads to dysfunctions beyond the immune system.

In humans, an intronic SNV in *MKLN1* was associated with urinary potassium excretion in Korean adults and another intronic *MKLN1* SNV with early bipolar disorder [31,32]. Furthermore, *MKLN1* has been associated with asthma in independent GWASs. A SNV in *MKLN1* ranked among the top 100 SNVs associated with childhood asthma in a study sample of 429 affected-offspring trios from a European American population [33]. A different SNV in the 5'-UTR of *MKLN1* was associated with asthma in a population including patients with severe or difficult-to-treat asthma [34].

In the ExAC database, only one *MKLN1* missense, but no nonsense, frameshift or splice site variants present in a homozygous state were found [35,36]. Furthermore, the probability of loss of function (LoF, specified as nonsense, splice acceptor, and splice donor variants) tolerance was estimated to be 1.00, indicating that the *MKLN1* gene is extremely LoF intolerant [37]. Therefore, it is conceivable that loss of function variants on both alleles might lead to severe phenotypes in humans.

To our knowledge, no link between muskelin 1 and zinc or copper metabolism has been reported to date. While acrodermatitis enteropatica in humans and acrodermatitis in cattle clinically resemble LAD in dogs, these diseases may be caused by completely different

molecular mechanisms. The fact that findings on zinc levels in the few published studies on LAD affected dogs were contradictory and zinc supplementation did not lead to improvement of lesions [6] support this hypothesis.

In conclusion, we identified the *MKLN1*:c.400+3A>C variant leading to a splice defect in the *MKLN1* gene as candidate causative variant for LAD in Bull Terriers and Miniature Bull Terriers. The molecular pathogenesis of LAD remains unclear. Our data facilitate genetic testing of Bull Terriers and Miniature Bull Terriers to prevent the unintentional breeding of LAD affected dogs. LAD affected dogs may serve as models to further clarify the elusive physiological role of muskelin 1 *in vivo*.

Materials and methods

Ethics statement

All animal experiments were performed according to the local regulations. The dogs in this study were examined with the consent of their owners. The study was approved by the "Cantonal Committee For Animal Experiments" (Canton of Bern; permits 22/07, 23/10, and 75/16).

Animals and samples

Bull Terriers with their characteristic egg-shaped head were founded as a dog breed in the 1850s in the United Kingdom. Originally, there were no size standards in this breed and smaller dogs were bred as a variety of the regular Bull Terrier. Eventually, two sub-populations formed and the Miniature Bull Terrier with a maximum height of 35.5 cm was recognized as an independent breed in 1991 by the American Kennel Club (AKC) and in 2011 by the European Fédération Cynologique Internationale (FCI). Therefore, Bull Terriers and Miniature Bull Terriers share a common ancestral gene pool, but represent independent closed populations today.

This study included samples from 251 Bull Terriers (41 LAD cases / 210 controls) and 89 Miniature Bull Terriers (5 LAD cases / 84 controls). Case/controls status was based on owners' reports. We additionally used 462 dogs from 62 breeds, which were assumed to be free of the disease allele (S3 Table). Skin biopsies were taken from two LAD affected Bull Terriers from toe, nose, lip, and forearm and fixed in 10% buffered formalin for 24 hours. Biopsies were processed, embedded in paraffin and sectioned at 4 μm. Skin sections were stained with hematoxylin and eosin. The histopathology was performed by veterinary pathologists (BR, Dipl.-ECVP, and SH). Two further biopsies from comparable sites of the same dogs were submerged in RNAlater solution for subsequent RNA isolation.

DNA isolation and SNV genotyping

We isolated genomic DNA from EDTA blood samples. Seventy-eight dogs were genotyped for either 173,662 or 218,256 SNVs on the illumina canine_HD chip. The raw SNV genotypes are available at https://www.animalgenome.org/repository/pub/BERN2017.1208/.

GWAS

The initial dataset consisted of 78 dogs and 220,853 markers. Using Plink version 1.9 [38] we excluded markers that were not located on autosomes or the X chromosome (n = 2,327) and markers with a genotyping rate lower than 90% (n = 49,814). Using the R package GenABEL [39] and the command "check.markers", dogs with a call rate < 90% (n = 3), ibs > 95% (n = 0), high individual heterozygosity (FDR = 0.01) (n = 1, included in dogs with low call rate) as well as markers with a maf < 1% (n = 55,931) and a genotyping rate < 90% (n = 10,951) were

excluded. Five outliers in the multidimensional scaling plot based on a genomic distance matrix were also removed. In a second quality control step, markers deviating from Hardy-Weinberg equilibrium (FDR = 0.2) in controls (n = 1,239), markers with a genotyping rate < 90% (n = 0) and maf < 1% (n = 26,215) were excluded, resulting in a final dataset of 70 dogs (22 cases, 48 controls) and 76,419 markers. A polygenic model of the hglm package [40], with a kinship matrix based on autosomal markers in the cleaned dataset as random effect, was estimated and a score test for association using the function "mmscore" was performed. The genomic inflation factor was 1.16. We corrected for multiple testing using Bonferroni correction with a significance level of 0.05. QQ plots were created using qqman version 0.1.4 [41].

Haplotype analysis

We visually inspected plink tped files for the region of interest on chromosome 14 using Excel and searched for homozygous regions with haplotype sharing in cases with a call rate >90%. The first flanking heterozygous markers on either side of the homozygous region in 22 cases defined the borders of the critical interval.

Whole genome sequencing of an affected Bull Terrier

An Illumina PCR-free TruSeq fragment library with 350 bp insert size of an LAD affected Bull Terrier was prepared. We collected 219 million 2 x 150 bp paired-end reads or 24x coverage on a HiSeq3000 instrument. The reads were mapped to the dog reference genome assembly CanFam3.1 and aligned using Burrows-Wheeler Aligner (BWA) version 0.7.5a [42] with default settings. The generated SAM file was converted to a BAM file and the reads were sorted by coordinate using samtools [43]. Picard tools (http://sourceforge.net/projects/picard/) was used to mark PCR duplicates. To perform local realignments and to produce a cleaned BAM file, we used the Genome Analysis Tool Kit (GATK version 2.4.9, 50) [44]. GATK was also used for base quality recalibration with canine dbsnp version 139 data as training set. The sequence data were deposited under the study accession PRJEB16012 and sample accession SAMEA4504844 at the European Nucleotide Archive.

Variant calling

Putative SNVs were identified in each of 192 samples (S2 Table) individually using GATK HaplotypeCaller in gVCF mode [45]. Subsequently all sample gVCF files were joined using Broad GenotypeGVCFs walker (-stand_emit_conf 20.0; -stand_call_conf 30.0). Filtering was performed using the variant filtration module of GATK using the following standard filters: SNVs: Quality by Depth: QD < 2.0; Mapping quality: MQ < 40.0; Strand filter: FS > 60.0; MappingQualityRankSum: MQRankSum < -12.5; ReadPosRankSum < -8.0. INDELs: Quality by Depth: QD < 2.0; Strand filter: FS > 200.0. The functional effects of the called variants were predicted using SnpEFF software [46] together with the NCBI annotation release 104 on CanFam 3.1. For the filtering of candidate causative variants in the case, we used 191 control genomes, which were either publicly available [47] or produced during other projects of our group or contributed by members of the Dog Biomedical Variant Database Consortium. A detailed list of these control genomes is given in S2 Table.

Gene analysis

We used the dog CanFam 3.1 reference genome assembly for all analyses. Numbering within the canine *MKLN1* gene corresponds to the accessions XM_005628367.3 (mRNA) and

XP_005628424.1 (protein). Numbering within the human *MKLN1* gene corresponds to the accessions NM_013255.4 (mRNA) and NP_037387.2 (protein).

Sanger sequencing

We used Sanger sequencing to confirm the candidate variant *MKLN1*:c.400+3A>C and to genotype the dogs in this study. A 797 bp fragment containing the variant was PCR amplified from genomic DNA using AmpliTaq Gold 360 Master Mix (Life Technologies) and the primers CCATGCACTGTAGCCACATC and TGGAAAAGGTTCCACTTGAAAT. After treatment with shrimp alkaline phosphatase and endonuclease I, PCR products were directly sequenced on an ABI 3730 capillary sequencer (Life Technologies). We analyzed the Sanger sequence data using the software Sequencher 5.1 (GeneCodes).

RNA isolation and RT-PCR

RNA was extracted from skin samples using the RNeasy Fibrous Tissue Mini Kit (Qiagen). The tissue was first finely crushed by mechanical means using TissueLyser (Quiagen), and RNA was extracted by centrifugation following the instructions by the manufacturer. Total mRNA was reverse transcribed into cDNA using the SuperScript IV Reverse Transcriptase kit (Thermo Fisher) with oligo d(T) primers. A PCR on the synthesized cDNA was carried out using primer MKLN1_c_F2, CCTCCCCAGTACTTGATTCTG, located at the boundary of exons 2 and 3, and primer MKLN1_c_R2, TTCCTGTTCACGGTACTTGC, located at the boundary of exons 5 and 6 of the *MKLN1* gene. The products were analyzed on a Fragment Analyzer capillary gel electrophoresis instrument (Advanced Analytical). The sequence of the obtained RT-PCR products was confirmed by Sanger sequencing as described above.

Supporting information

S1 Fig. Sequence context of the MKLN1 variant.
(PDF)

S2 Fig. Predicted amino acid sequences of the wildtype and mutant MKLN1 proteins.
(PDF)

S1 Table. Private variants in the critical interval.
(XLSX)

S2 Table. Control dogs used for whole genome sequencing.
(XLSX)

S3 Table. Control dogs genotyped for the MKLN1 variant.
(XLSX)

Acknowledgments

The authors are grateful to the dog owners who donated samples and shared pedigree data and phenotype information of their dogs. Sample collections were supported through various programs. French samples were collected by a network of veterinarians and Laetitia Lagoutte and Nadine Botherel managing the Cani-DNA biobank (http://dog-genetics.genouest.org), which is part of the CRB-Anim infrastructure (ANR-11-INBS-0003, in the framework of the "Investing for the Future" program (PIA)). We thank Nathalie Besuchet Schmutz, Muriel Fragnière, Kaisu Hiltunen, Fabiana Jakob, and Sabrina Schenk for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing the high-

throughput sequencing experiments, and the Interfaculty Bioinformatics Unit of the University of Bern for providing high performance computing infrastructure. We thank the Dog Biomedical Variant Database Consortium (Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Cord Drögemüller, Oliver Forman, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Vidhya Jagannathan, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Anita Oberbauer, Jeffrey Schoenebeck, Sheila Schmutz, Claire Wade) for sharing whole genome sequencing data from control dogs. We also acknowledge all canine researchers who deposited dog whole genome sequencing data into public databases.

Author Contributions

Conceptualization: Tosso Leeb.

Data curation: Vidhya Jagannathan.

Formal analysis: Anina Bauer, Vidhya Jagannathan, Tosso Leeb.

Funding acquisition: Catherine André, Hannes Lohi, Monika M. Welle, Petra Roosje, Tosso Leeb.

Investigation: Anina Bauer, Vidhya Jagannathan, Sandra Högler, Barbara Richter, Marjo K. Hytönen, Cathryn Mellersh, Margret L. Casal, Tosso Leeb.

Methodology: Anina Bauer.

Project administration: Tosso Leeb.

Resources: Neil A. McEwan, Anne Thomas, Edouard Cadieu, Catherine André, Marjo K. Hytönen, Hannes Lohi, Monika M. Welle, Petra Roosje, Cathryn Mellersh, Margret L. Casal.

Supervision: Tosso Leeb.

Visualization: Tosso Leeb.

Writing – original draft: Anina Bauer, Tosso Leeb.

Writing – review & editing: Anina Bauer, Vidhya Jagannathan, Sandra Högler, Barbara Richter, Neil A. McEwan, Anne Thomas, Edouard Cadieu, Catherine André, Marjo K. Hytönen, Hannes Lohi, Monika M. Welle, Petra Roosje, Cathryn Mellersh, Margret L. Casal, Tosso Leeb.

References

1. Moynahan EJ (1974) Acrodermatitis enteropathica: a lethal inherited human zinc-deficiency disorder. Lancet 2: 399–400. PMID: 4136854
2. Moynahan EJ (1982) The Lancet: Acrodermatitis Enteropathica: A Lethal Inherited Human Zinc-Deficiency Disorder. Nutr Rev 40: 84–86. PMID: 7050775
3. Lakdawala N, Grant-Kels JM (2015) Acrodermatitis enteropathica and other nutritional diseases of the folds (intertriginous areas). Clin Dermatol 33: 414–419. https://doi.org/10.1016/j.clindermatol.2015.04.002 PMID: 26051055
4. Küry S, Dréno B, Bézieau S, Giraudet S, Kharfi M, Kamoun R, et al. (2002) Identification of *SLC39A4*, a gene involved in acrodermatitis enteropathica. Nat Genet 31: 239–240. https://doi.org/10.1038/ng913 PMID: 12068297
5. Yuzbasiyan-Gurkan V, Bartlett E (2006) Identification of a unique splice site variant in SLC39A4 in bovine hereditary zinc deficiency, lethal trait A46: An animal model of acrodermatitis enteropathica. Genomics 88: 521–526. https://doi.org/10.1016/j.ygeno.2006.03.018 PMID: 16714095
6. Jezyk PF, Haskins ME, MacKay-Smith WE, Patterson DF (1986) Lethal acrodermatitis in bull terriers. J Am Vet Med Assoc 188: 833–839. PMID: 3710872

7. McEwan NA (1990) Lethal acrodermatitis of bull terriers. Vet Rec 127: 95.
8. McEwan NA, McNeil PE, Thompson H, McCandlish IA (2000) Diagnostic features, confirmation and disease progression in 28 cases of lethal acrodermatitis of bull terriers. J Small Anim Pract 41: 501–507. PMID: 11105789
9. McEwan NA (2001) Malassezia and Candida infections in bull terriers with lethal acrodermatitis. J Small Anim Pract 42: 291–297. PMID: 11440398
10. McEwan NA, Huang HP, Mellor DJ (2003) Immunoglobulin levels in Bull terriers suffering from lethal acrodermatitis. Vet Immunol Immunopathol 96: 235–238. PMID: 14592736
11. Uchida Y, Moon-Fanelli AA, Dodman NH, Clegg MS, Keen CL (1997) Serum concentrations of zinc and copper in bull terriers with lethal acrodermatitis and tail-chasing behavior. Am J Vet Res 58: 808–810. PMID: 9256960
12. Grider A, Mouat MF, Mauldin EA, Casal ML (2007) Analysis of the liver soluble proteome from bull terriers affected with inherited lethal acrodermatitis. Mol Genet Metab 92: 249–257. https://doi.org/10.1016/j.ymgme.2007.07.003 PMID: 17693109
13. Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R (2006) Comprehensive splice-site analysis using comparative genomics. Nucl Acids Res 34: 3955–3967. https://doi.org/10.1093/nar/gkl556 PMID: 16914448
14. Ravi Lab SpliceRack, Motif Frequency, available from: http://katahdin.mssm.edu/splice/viewsplicemotifgraphform.cgi?database = spliceNew
15. Martoni E, Urciuolo A, Sabatelli P, Fabris M, Bovolenta M, Neri M, et al. (2009) Identification and characterization of novel collagen VI non-canonical splicing mutations causing Ullrich congenital muscular dystrophy. Hum Mutat 30: E662–E672. https://doi.org/10.1002/humu.21022 PMID: 19309692
16. Kawarai T, Montecchiani C, Miyamoto R, Gaudiello F, Caltagirone C, Izumi Y, et al. (2017) Spastic paraplegia type 4: A novel *SPAST* splice site donor mutation and expansion of the phenotype variability. J Neurol Sci 380: 92–97. https://doi.org/10.1016/j.jns.2017.07.011 PMID: 28870597
17. Koppolu AA, Madej-Pilarczyk A, Rydzanicz M, Kosińska J, Gasperowicz P, Dorszewska J, et al. (2017) A novel *de novo COL6A1* mutation emphasizes the role of intron 14 donor splice site defects as a cause of moderate-progressive form of ColVI myopathy - a case report and review of the genotype-phenotype correlation. Folia Neuropathol 55: 214–220. https://doi.org/10.5114/fn.2017.70486 PMID: 28984114
18. Adams JC, Seed B, Lawler J (1998) Muskelin, a novel intracellular mediator of cell adhesive and cytoskeletal responses to thrombospondin-1. EMBO 17: 4964–4974.
19. Francis O, Han F, Adams JC (2013) Molecular phylogeny of a RING E3 ubiquitin ligase, conserved in eukaryotic cells and dominated by homologous components, the muskelin/RanBPM/CTLH complex. PLoS One 8: e75217. https://doi.org/10.1371/journal.pone.0075217 PMID: 24143168
20. Prag S, Collett GDM, Adams JC (2004) Molecular analysis of muskelin identifies a conserved discoidin-like domain that contributes to protein self-association. Biochem J 381: 547–559. https://doi.org/10.1042/BJ20040253 PMID: 15084145
21. Delto CF, Heisler FF, Kuper J, Sander B, Kneussel M, Schindelin H (2015) The LisH motif of muskelin is crucial for oligomerization and governs intracellular localization. Structure 23: 364–373. https://doi.org/10.1016/j.str.2014.11.016 PMID: 25579817
22. Hasegawa H, Katoh H, Fujita H, Mori K, Negishi M (2000) Receptor isoform-specific interaction of prostaglandin EP3 receptor with muskelin. Biochem Biophys Res Commun 276: 350–354. https://doi.org/10.1006/bbrc.2000.3467 PMID: 11006128
23. Gueron G, Giudice J, Valacco P, Paez A, Elguero B, Toscani M, et al. (2014) Heme-oxygenase-1 implications in cell morphology and the adhesive behavior of prostate cancer cells. Oncotarget 5: 4087–4102. https://doi.org/10.18632/oncotarget.1826 PMID: 24961479
24. Umeda M, Nishitani H, Nishimoto T (2003). A novel nuclear protein, Twa1, and Muskelin comprise a complex with RanBPM. Gene 303: 47–54. PMID: 12559565
25. Kobayashi N, Yang J, Ueda A, Suzuki T, Tomaru K, Takeno M, et al. (2007) RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8α and ARMC8β are components of the CTLH complex. Gene 396: 236–247. https://doi.org/10.1016/j.gene.2007.02.032 PMID: 17467196
26. Valiyaveettil M, Bentley AA, Gursahaney P, Hussien R, Chakravarti R, Kureishy N, et al. (2008) Novel role of the muskelin-RanBP9 complex as a nucleocytoplasmic mediator of cell morphology regulation. J Cell Biol 182: 727–739. https://doi.org/10.1083/jcb.200801133 PMID: 18710924
27. Debenedittis P, Harmelink C, Chen Y, Wang Q, Jiao K (2011) Characterization of the novel interaction between muskelin and TBX20, a critical cardiogenic transcription factor. Biochem Biophys Res Commun 409: 338–343. https://doi.org/10.1016/j.bbrc.2011.05.020 PMID: 21586270

28. Ledee DR, Gao CY, Seth R, Fariss RN, Tripathi BK, Zelenka PS (2005) A specific interaction between muskelin and the cyclin-dependent kinase 5 activator p39 promotes peripheral localization of muskelin. J Biol Chem 280: 21376–21383. https://doi.org/10.1074/jbc.M501215200 PMID: 15797862

29. Tripathi BK, Lowy DR, Zelenka PS (2015) The Cdk5 activator P39 specifically links muskelin to myosin II and regulates stress fiber formation and actin organization in lens. Exp Cell Res 330: 186–198. https://doi.org/10.1016/j.yexcr.2014.08.003 PMID: 25128817

30. Heisler FF, Loebrich S, Pechmann Y, Maier N, Zivkovic AR, Tokito M, et al. (2011) Muskelin regulates actin filament- and microtubule-based GABA(A) receptor transport in neurons. Neuron 70: 66–81. https://doi.org/10.1016/j.neuron.2011.03.008 PMID: 21482357

31. Park YM, Kwock CK, Kim K, Kim J, Yang YJ (2017) Interaction between single nucleotide polymorphism and urinary sodium, potassium, and sodium-potassium ratio on the risk of hypertension in Korean adults. Nutrients 9: E235. https://doi.org/10.3390/nu9030235 PMID: 28273873

32. Nassan M, Li Q, Croarkin PE, Chen W, Colby CL, Veldic M, et al. (2017) A genome wide association study suggests the association of muskelin with early onset bipolar disorder: Implications for a GABAergic epileptogenic neurogenesis model. J Affect Disord 208: 120–129. https://doi.org/10.1016/j.jad.2016.09.049 PMID: 27769005

33. Ding L, Abebe T, Beyene J, Wilke RA, Goldberg A, Woo JG, et al. (2013) Rank-based genome-wide analysis reveals the association of Ryanodine receptor-2 gene variants with childhood asthma among human populations. Hum Genomics 7: 16. https://doi.org/10.1186/1479-7364-7-16 PMID: 23829686

34. 'Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, et al. (2010) Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. J Allergy Clin Immunol 125: 328–335.e11. https://doi.org/10.1016/j.jaci.2009.11.018 PMID: 20159242

35. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. (2017) The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res 45: D840–D845. https://doi.org/10.1093/nar/gkw971 PMID: 27899611

36. ExAC Browser (Beta), Exome Aggregation Consortium, available from http://exac.broadinstitute.org/

37. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536: 285–291. https://doi.org/10.1038/nature19057 PMID: 27535533

38. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575. https://doi.org/10.1086/519795 PMID: 17701901

39. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM (2007) GenABEL: an R library for genome-wide association analysis. Bioinformatics 23: 1294–1296. https://doi.org/10.1093/bioinformatics/btm108 PMID: 17384015

40. Ronnegard L, Shen X, Alam M (2010) hglm: A package for fitting hierarchical generalized linear models. The R Journal 2: 20–28.

41. Turner S (2017) qqman: Q-Q and Manhattan Plots for GWAS Data. R package version 0.1.4. https://CRAN.R-project.org/package = qqman

42. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168

43. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987–2993. https://doi.org/10.1093/bioinformatics/btr509 PMID: 21903627

44. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297–1303. https://doi.org/10.1101/gr.107524.110 PMID: 20644199

45. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 43: 11.0.1–33. https://doi.org/10.1002/0471250953.bi1110s43 PMID: 25431634

46. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6: 80–92. https://doi.org/10.4161/fly.19695 PMID: 22728672

47. Bai B, Zhao WM, Tang BX, Wang YQ, Wang L, Zhang Z, et al. (2015) DoGSD: the dog and wolf genome SNP database. Nucleic Acids Res 43 (Database issue): D777–783. https://doi.org/10.1093/nar/gku1174 PMID: 25404132

A splice site variant in the *SUV39H2* gene in Greyhounds with nasal parakeratosis

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	genetic analyses, original draft, review and editing of manuscript, illustrations

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12643

A splice site variant in the *SUV39H2* gene in Greyhounds with nasal parakeratosis

A. Bauer*†, J. Nimmo‡, R. Newman§, M. Brunner†¶, M. M. Welle†¶, V. Jagannathan*† and T. Leeb*†

*Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland. †DermFocus, University of Bern, 3001 Bern, Switzerland. ‡ASAP Laboratory, Mulgrave, Vic. 3170, Australia. §Mobile Vet Services and Supplies, Warwick, Qld 4370, Australia. ¶Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland.

Summary

Hereditary nasal parakeratosis (HNPK), described in the Labrador Retriever breed, is a monogenic autosomal recessive disorder that causes crusts and fissures on the nasal planum of otherwise healthy dogs. Our group previously showed that this genodermatosis may be caused by a missense variant located in the *SUV39H2* gene encoding a histone 3 lysine 9 methyltransferase, a chromatin modifying enzyme with a potential role in keratinocyte differentiation. In the present study, we investigated a litter of Greyhounds in which six out of eight puppies were affected with parakeratotic lesions restricted to the nasal planum. Clinically and histologically, the lesions were comparable to HNPK in Labrador Retrievers. Whole genome sequencing of one affected Greyhound revealed a 4-bp deletion at the 5′-end of intron 4 of the *SUV39H2* gene that was absent in 188 control dog and three wolf genomes. The variant was predicted to disrupt the 5′-splice site with subsequent loss of SUV39H2 function. The six affected puppies were homozygous for the variant, whereas the two non-affected littermates were heterozygous. Genotyping of a larger cohort of Greyhounds revealed that the variant is segregating in the breed and that this breed might benefit from genetic testing to avoid carrier × carrier matings.

Keywords *Canis lupus* familiaris, chromatin, differentiation, dog, epigenetics, genodermatosis, hereditary nasal parakeratosis, keratinocyte, skin, whole genome sequencing

The epidermis, the outermost layer of the skin, undergoes a continuous process of self-renewal throughout a lifetime. Keratinocytes migrate from the basal layer to the stratum corneum as they differentiate. Terminally differentiated keratinocytes are present as corneocytes without nuclei and organelles that form the tightly sealing outermost layer of the epidermis before being sloughed from the skin surface (Candi *et al.* 2005; Fuchs 2007). The terminal differentiation of keratinocytes is tightly controlled, and there is increasing evidence that epigenetic processes play an important role in keratinocyte differentiation (Botchkarev *et al.* 2012). Variants in genes encoding structural components, cell cycle regulators and adhesion molecules involved in this process can lead to heritable skin diseases, so called genodermatoses.

Address for correspondence

T. Leeb, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland.
E-mail: tosso.leeb@vetsuisse.unibe.ch

Accepted for publication 19 December 2017

Our group previously identified the *SUV39H2*: XM_535179.5:c.972T>G; p.Asn324Lys missense variant to be causative for hereditary nasal parakeratosis (HNPK) in Labrador Retrievers (Jagannathan *et al.* 2013). This monogenic autosomal recessive skin disorder leads to crusts and fissuring of the nasal planum and has so far been described only in the Labrador Retriever breed (Pagé *et al.* 2003; Peters *et al.* 2003). *SUV39H2* encodes the suppressor of variegation 3–9 homolog 2 (Drosophila), a histone 3 lysine 9 (H3K9) methyltransferase which mediates chromatin silencing (O'Carroll *et al.* 2000; Jenuwein & Allis 2001; Peters *et al.* 2001). The *SUV39H2* missense variant detected in Labrador Retriever HNPK affects the catalytic site and abolishes enzyme function (Schuhmacher *et al.* 2015). Loss of *SUV39H2* function may result in delayed terminal keratinocyte differentiation, which seems a plausible pathomechanism to explain the HNPK phenotype in Labrador Retrievers (Jagannathan *et al.* 2013).

In the present study, we investigated a Greyhound litter with eight puppies. Six puppies were affected with nasal parakeratosis to varying degrees of severity (Fig. 1). Only

the nasal planum was affected; the dogs' other skin appeared normal. Topical treatment with zinc cream led to improvement of the lesions in one case but not in others. Clinically and histologically, the lesions were comparable to the changes observed in Labrador Retrievers with HNPK. The parents were not affected, and genetic testing of the mother revealed that she was not a carrier for the *SUV39H2*:c.972T>G variant causing HNPK in Labrador Retrievers (Orivet Genetics). Based on these data, we hypothesized that a novel genetic variant, inherited in a monogenic autosomal recessive mode, was responsible for the nasal parakeratosis in the investigated Greyhounds. We isolated DNA from EDTA blood from all eight puppies and submitted them for genotyping on the canine Illumina HD 220 K SNP chip (Neogen/GeneSeek). The SNP genotypes are publicly available at https://www.animalgenome.org/repository/pub/BERN2017.1102/. We used these genotype data and PLINK version 1.07 (Purcell *et al.* 2007) to search for extended regions of homozygosity with allele sharing equal to 1 Mb or more present in all six cases. Markers located on the sex chromosomes were excluded. Using default settings, this resulted in 48 homozygous regions totalling 122 Mb or 5% of the canine genome. The largest homozygous segment spanned ~22 Mb, was located on chromosome 2 and harboured the *SUV39H2* gene (Table S1).

We prepared a PCR-free DNA library of an affected Greyhound, collected 2 × 150-bp reads on an Illumina HiSeq 3000 instrument and re-sequenced the genome at 37× coverage (ENA project accession no. PRJEB16012, sample accession no. SAMEA104125118). Variants were called with respect to the CanFam 3.1 reference genome assembly, and

Figure 1 Nasal parakeratosis in two male Greyhounds of the same litter. Note the varying severity of the lesions affecting exclusively the nasal planum. The affected dogs did not show any other clinical signs than the changes of the nose.

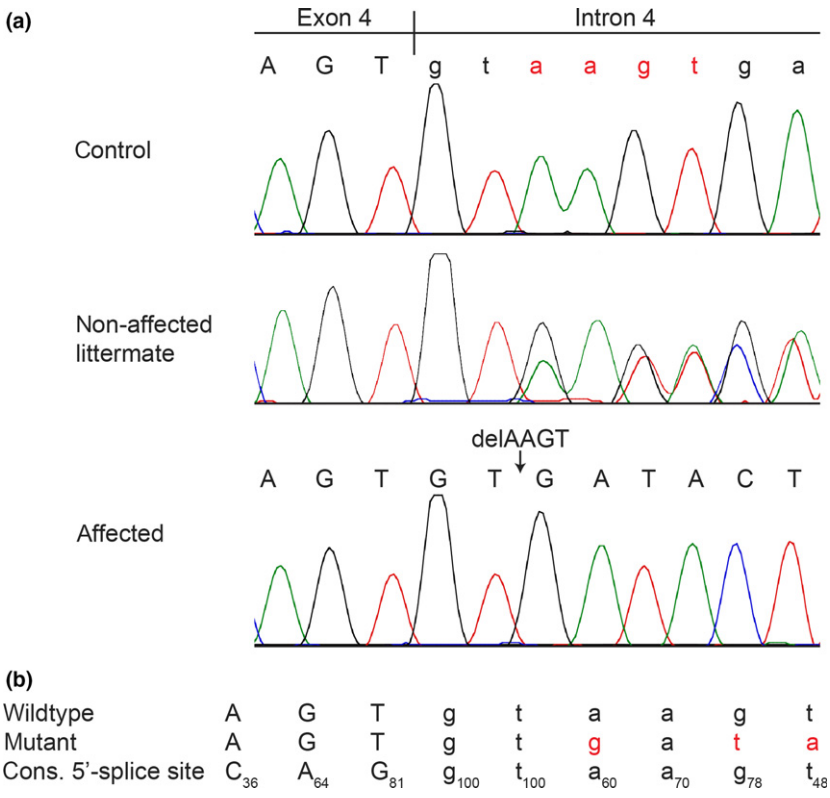


Figure 2 Electropherograms showing the *SUV39H2*:c.996+3_996+6delAAGT variant and its impact on the splice donor motif. (a) Electropherograms showing the wildtype splice site in a control dog, a non-affected littermate with a wt/del genotype and an affected Greyhound homozygous for the deletion. The four nucleotides deleted in the mutant allele are shown in red in the wildtype sequence. (b) Wildtype and mutant allele compared to the consensus sequence for the human U2 GT-AG type 5′-splice sites (Sheth *et al.* 2006). Subscript numbers in the consensus sequence indicate the percentage of the respective conserved nucleotide in 183 682 investigated human 5′-splice site motifs of the U2 GT-AG type. Note that the wildtype sequence already deviates from the perfect consensus sequence at three of the nine positions. The mutant sequence differs at six of the nine nucleotides in the U1 spliceosomal RNA recognition site. The additional differences to the optimal consensus are highlighted in red.

they were compared to 188 control dog and three wolf genomes as described previously (Table S2; Bauer *et al.* 2017). We filtered for variants that were present in a homozygous state in the affected Greyhound and absent in the control genomes, assuming that the causative variant was present only in the Greyhound breed. We detected 28 variants with high, moderate or low predicted impact on protein function (Table S3). Eight were located within a shared homozygous region in the cases including a splice site variant in the *SUV39H2* gene, our primary functional candidate gene for nasal parakeratosis. This variant was a 4-bp deletion within the 5′-splice site of intron 4, XM_535179.6:c.996+3_996+6delAAGT or Chr2:21,731,812_21,731,815delACTT (CanFam 3.1).

The mutant sequence at the 5′-splice site retained the strictly conserved GT-dinucleotide at the first two bases of the intron. However, it differed in six out of nine nucleotides of the entire 5′-splice site (Fig. 2). To assess the likelihood of a splice defect, we analysed the frequency of the wildtype and mutant sequence motifs in a compilation of 186 630 human 5′-splice sites (http://katahdin.mssm.edu/splice/viewsplicemotifgraphform.cgi?database=spliceNew; Sheth *et al.* 2006). The canine wildtype sequence AGTgtaagt was identical to the sequence of 288 human 5′-splice sites, whereas the mutant sequence motif AGTgtgata did not occur in human 5′-splice sites. Splice site prediction software also clearly recognized the wildtype sequence but not the mutant sequence (http://www.fruitfly.org/seq_tools/splice.html; Reese *et al.* 1997). Thus, both *in silico* analyses suggested that the mutant 5′-splice site in Greyhounds with nasal parakeratosis is non-functional.

As we did not have access to RNA from an affected dog, we experimentally assessed the functional consequence on splicing by an RNA-seq experiment on skin RNA from a heterozygous Greyhound in comparison to a homozygous wildtype Greyhound (project accession no. PRJEB21761, sample accessions nos. SAMEA104393648 and SAMEA104393651). This experiment demonstrated retention of intron 4 in transcripts originating from the mutant allele in the heterozygous dog (Fig. S1).

We genotyped the eight puppies for this variant by Sanger sequencing and found a perfect association with the phenotype: The six affected puppies were homozygous for the deletion, and the two non-affected littermates were both heterozygous carriers. We then genotyped a larger cohort of Greyhounds as well as 483 control dogs of different breeds for this variant (Table S4). As expected, the variant was absent from all tested breeds other than the Greyhounds. In our cohort of 420 Greyhounds, which were not closely related to the investigated litter, we did not detect any homozygous mutant dogs but found eight additional heterozygous dogs. These data indicate a carrier frequency of roughly 2% in the Greyhound population.

In light of the previous knowledge about the *SUV39H2*: c.972C>T variant in Labrador Retrievers with HNPK, we think that our data strongly suggest that *SUV39H2*: c.996+3_996+6delAAGT is the causative genetic variant underlying the observed nasal parakeratosis in Greyhounds. The Greyhound breed might benefit from genetic testing to avoid future carrier × carrier matings.

Acknowledgements

The authors would like to thank the dog owners for donating samples and pictures and for sharing information of their dogs. The authors also wish to thank Nathalie Besuchet Schmutz, Muriel Fragnière and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure. We acknowledge collaborators of the Dog Biomedical Variant Database Consortium (DBVDC), Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Cord Drögemüller, Kari Ekenstedt, Kiterie Faller, Oliver Forman, Steve Friedenberg, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Leonardo Murgiano, Anita Oberbauer, Sheila Schmutz, Jeffrey Schoenebeck, Kim Summers, Frank van Steenbeek and Claire Wade for sharing dog genome sequence data from control dogs and wolves. This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738/1) and the Albert-Heim Foundation (no. 105).

References

Bauer A., Waluk D.P., Galichet A. *et al.* (2017) A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. *PLoS Genetics* **13**, e1006651.

Botchkarev V.A., Gdula M.R., Mardaryev A.N., Sharov A.A. & Fessing M.Y. (2012) Epigenetic regulation of gene expression in keratinocytes. *Journal of Investigative Dermatology* **132**, 2505–21.

Candi E., Schmidt R. & Melino G. (2005) The cornified envelope: a model of cell death in the skin. *Nature Reviews Molecular Cell Biology* **6**, 328–40.

Fuchs E. (2007) Scratching the surface of skin development. *Nature* **445**, 834–42.

Jagannathan V., Bannoehr J., Plattet P. *et al.* (2013) A mutation in the *SUV39H2* gene in Labrador Retrievers with hereditary nasal parakeratosis (HNPK) provides insights into the epigenetics of keratinocyte differentiation. *PLoS Genetics* **9**, e1003848.

Jenuwein T. & Allis C.D. (2001) Translating the histone code. *Science* **293**, 1074–80.

O'Carroll D., Scherthan H., Peters A.H. *et al.* (2000) Isolation and characterization of *Suv39 h2*, a second histone H3 methyltransferase gene that displays testis-specific expression. *Molecular and Cellular Biology* **20**, 9423–33.

Pagé N., Paradis M., Lapointe J.M. & Dunstan R.W. (2003) Hereditary nasal parakeratosis in Labrador Retrievers. *Veterinary Dermatology* **14**, 103–10.

Peters A.H., O'Carroll D., Scherthan H. *et al.* (2001) Loss of the Suv39 h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–37.

Peters J., Scott D.W., Erb H.N. & Miller W.H. (2003) Hereditary nasal parakeratosis in Labrador retrievers: 11 new cases and a retrospective study on the presence of accumulations of serum ('serum lakes') in the epidermis of parakeratotic dermatoses and inflamed nasal plana of dogs. *Veterinary Dermatology* **14**, 197–203.

Purcell S., Neale B., Todd-Brown K. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* **81**, 559–75.

Reese M.G., Eeckman F.H., Kulp D. & Haussler D. (1997) Improved splice site detection in Genie. *Journal of Computational Biology* **4**, 311–23.

Schuhmacher M.K., Kudithipudi S., Kusevic D., Weirich S. & Jeltsch A. (2015) Activity and specificity of the human SUV39H2 protein lysine methyltransferase. *Biochimica et Biophysica Acta* **1849**, 55–63.

Sheth N., Roca X., Hastings M.L., Roeder T., Krainer A.R. & Sachidanandam R. (2006) Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Research* **34**, 3955–67.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1. *SUV39H2* splice defect.

Table S1. Homozygous intervals ≥1 Mb shared between six cases.

Table S2. Information on 192 dog/wolf genome sequences.

Table S3. Private variants detected by whole genome sequencing.

Table S4. Genotypes of 420 Greyhounds and 483 control dogs from 65 various other breeds.

A single base deletion in the *SLC45A2* gene in a Bullmastiff with oculocutaneous albinism

Journal:	Animal Genetics
Manuscript status:	published
Contributions:	supervision of genetic analyses, review and editing of manuscript

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

SHORT COMMUNICATION

doi: 10.1111/age.12582

A single base deletion in the *SLC45A2* gene in a Bullmastiff with oculocutaneous albinism

M. Caduff*†[1], A. Bauer*†[1], V. Jagannathan*† and T. Leeb*†

*Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland. †DermFocus, University of Bern, 3001 Bern, Switzerland.

Summary

Oculocutaneous albinism type 4 (OCA4) in humans and similar phenotypes in many animal species are caused by variants in the *SLC45A2* gene, encoding a putative sugar transporter. In dog, two independent *SLC45A2* variants are known that cause oculocutaneous albinism in Doberman Pinschers and several small dog breeds respectively. For the present study, we investigated a Bullmastiff with oculocutaneous albinism. The affected dog was highly inbred and resulted from the mating of a sire to its own grandmother. We obtained whole genome sequence data from the affected dog and searched specifically for variants in candidate genes known to cause albinism. We detected a single base deletion in exon 6 of the *SLC45A2* gene (NM_001037947.1:c.1287delC) that has not been reported thus far. This deletion is predicted to result in an early premature stop codon. It was confirmed by Sanger sequencing and perfectly co-segregated with the phenotype in the available family members. We genotyped 174 unrelated dogs from diverse breeds, all of which were homozygous wildtype. We therefore suggest that *SLC45A2*:c.1287delC causes the observed oculocutaneous albinism in the affected Bullmastiff.

Keywords *Canis lupus familiaris*, coat colour, cream, dog, melanocyte, pigmentation

Coat colour in mammals is dependent on the synthesis of two different pigments in melanocytes: the black/brown eumelanin and the yellow/red phaeomelanin (Barsh 1996). Oculocutaneous albinism (OCA) in human comprises several autosomal recessive disorders for which this pigment synthesis is distorted, which results in a reduction of pigmentation in skin, hair and eyes. To date, seven different types, OCA 1–7, are known, and causative variants have been identified in six different genes, namely *TYR*, *OCA2*, *TYRP1*, *SLC45A2*, *SLC24A5* and *C10orf11* (Grønskov *et al.* 2007; Montoliu *et al.* 2014). Variants in *SLC45A2*, encoding solute carrier family 45 member 2, formerly also called *AIM1* or *MATP*, have been associated with OCA4 in humans (OMIM #606574; Newton *et al.* 2001). Underwhite mice, cream coloured horses, a single albinistic Western Lowland Gorilla, white tigers, silver chicken, albino zebrafish and orange-red coloured medaka b mutants were reported to be caused by variants in the *SLC45A2* gene (Fukamachi *et al.* 2001; Newton *et al.* 2001; Costin *et al.* 2003; Mariat *et al.* 2003; Gunnarsson *et al.* 2007; Dooley *et al.* 2013; Prado-Martinez *et al.* 2013; Xu *et al.* 2013). *SLC45A2* was investigated early on as a candidate gene in dog colour genetics, but initially no consistent variant explaining all cases of cream-coloured dogs could be identified (Schmutz & Berryere 2007). Subsequently, white Doberman Pinschers were shown to be homozygous for a large deletion of 4.1 kb including the last exon of *SLC45A2*. White Dobermans are characterized by OCA including health problems such as photophobia and an increased risk of cutaneous melanocytic neoplasms (Winkler *et al.* 2014). Several small, long-haired albinistic dogs, including a Lhasa Apso, a Pekinese, two Pomeranians and a mixed breed dog, were shown to be homozygous for a missense variant (p.Gly493Asp) in exon 7 of *SLC45A2*, leading to OCA (Wijesena & Schmutz 2015). Based on these findings, we hypothesized that a variant in *SLC45A2* could be responsible for the phenotype of the OCA-affected Bullmastiff in our study.

We received samples from an almost-white male Bullmastiff puppy and several of its relatives. The affected dog had only minimal pale pigmentation of the skin and hair and blue eyes (Fig. 1). The owner did not report any additional health problems such as eye squinting or

Address for correspondence

T. Leeb, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland.
E-mail: Tosso.Leeb@vetsuisse.unibe.ch

[1]These authors contributed equally to the study.

Accepted for publication 07 June 2017

Figure 1 Bullmastiff with oculocutaneous albinism.

photophobia. Pedigree analysis of the affected Bullmastiff revealed an inbreeding loop involving the mating of a male sire to his own grandmother. The OCA-affected dog was born in a litter of 10, from which eight puppies were euthanized after a few days due to atresia ani, a congenital anomaly of the rectum and anus (Fig. 2a). One of the two surviving puppies showed oculocutaneous albinism, and both parents and the other surviving puppy had a fawn phenotype, which is typical for Bullmastiffs.

The complete genome of the OCA-affected Bullmastiff was sequenced at 14× genome coverage using 2 × 150 bp reads on an Illumina HiSeq 3000 instrument (ENA project accession PRJEB16012; sample accession SAMEA103949042). Private variants were identified by comparing the obtained sequence to 373 control dog genomes generated for previous projects and the canine reference genome assembly CanFam 3.1 as described (Bauer *et al.* 2017). We detected 13 private homozygous protein-changing variants including a frameshifting single-base deletion in exon 6 of the *SLC45A2* gene (CanFam3.1:chr4:g.73 864 860delC; NM_001037947.1:c.1287delC; Table S1). This frameshift is predicted to result in a premature stop codon in the open reading frame (NP_001033036.1:p.(Met430CysfsTer4), which truncates the last 101 amino acids of the wildtype SLC45A2 protein.

We confirmed the variant in the affected Bullmastiff by Sanger sequencing (Fig. 2b). Primers SLC45A2_F1, CTGGTCCTTCCTTGTGGAAA and SLC45A2_R1, GAGCTG CAGAGGAAAGAGGA, were used to amplify exon 6 of the *SLC45A2* gene together with flanking sequences from the affected dog, five closely related family members as well as 174 unrelated dogs of various breeds with normal pigmentation (Table S2). Primer SLC45A2_F2, CACCGCCAGCTG TAATTTCT was used as the sequencing primer to obtain sequences on an ABI3730 capillary sequencer (Applied Biosystems). The Sanger sequencing confirmed perfect cosegregation of the recessive deletion allele with the phenotype in the investigated family. None of the 174 unrelated dogs from various breeds carried the mutant allele.

An interesting feature of *SLC45A2* variants is the genotype–phenotype correlation, considering the impact of the variant and its effect on the production of eumelanin versus phaeomelanin. The *SLC45A2* missense variant p.Ala477Val, with recessive inheritance in tigers, and two chicken missense variants, p.Tyr277Cys and p.Leu347Met with dominant inheritance, probably have the least impact and retain some functional activity (hypomorphic alleles). These alleles affect mostly phaeomelanin but not eumelanin production (Gunnarsson *et al.* 2007; Xu *et al.* 2013). In horses, the missense variant p.Asp153Asn prevents phaeomelanin synthesis in heterozygotes, whereas in homozygotes both phaeomelanin and eumelanin synthesis are inhibited (Mariat *et al.* 2003). Sex-linked imperfect albinism in Japanese quail, underwhite in mice and oculocutaneous albinism in Doberman Pinschers are caused by recessive null alleles involving frame-shift or deletion variants, where both eumelanin and phaeomelanin synthesis is blocked (Newton *et al.* 2001; Gunnarsson *et al.* 2007; Winkler *et al.* 2014). Although most *SLC45A2* missense variants in domestic animals have an effect primarily on phaeomelanin, p.Gly493Asp found in small dog breeds apparently has a more profound effect and also blocks phaeomelanin and eumelanin production (Wijesena & Schmutz 2015).

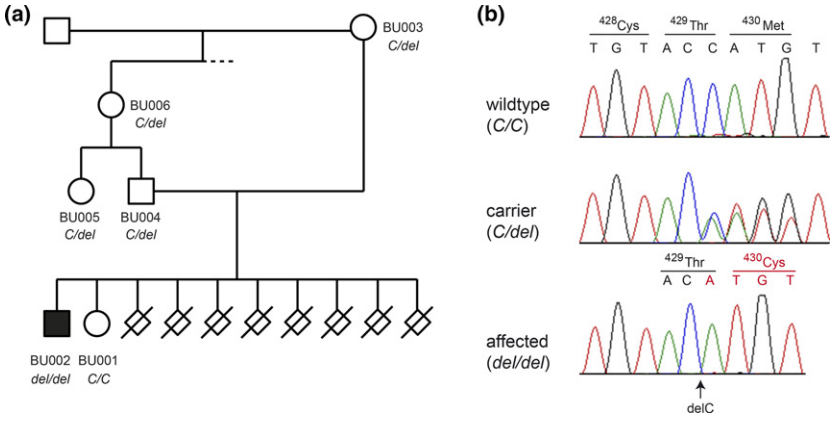


Figure 2 Genotypes of the affected Bullmastiff and its relatives. (a) The pedigree of the Bullmastiff family reveals an inbreeding loop. Genotypes of the sampled dogs are indicated. (b) Sanger sequencing chromatograms illustrate sequences of a homozygous wildtype dog, a heterozygous carrier and the homozygous affected dog. A homozygous deletion of a single cytosine is visible in the dog with oculocutaneous albinism (*SLC45A2:* c.1287delC).

In our study, we identified a new *SLC45A2* variant, which most likely causes oculocutaneous albinism in a Bullmastiff. The effect of this frame-shift variant is in agreement with previous results, as it leads to a nearly complete loss of both phaeomelanin and eumelanin production.

Acknowledgements

The authors would like to thank the dog breeder for donating samples and pictures and for sharing pedigree information of her dogs. The authors also wish to thank Nathalie Besuchet, Muriel Fragnière and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure. This study was supported by grants from the Swiss National Science Foundation (31003A_149313) and the Albert-Heim Foundation (no. 105).

References

Barsh G.S. (1996) The genetics of pigmentation: from fancy genes to complex traits. *Trends in Genetics* **12**, 299–305.

Bauer A., Waluk D.P., Galichet A. *et al.* (2017) A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. *PLoS Genetics* **13**, e1006651.

Costin G.E., Valencia J.C., Vieira W.D., Lamoreux M.L. & Hearing V.J. (2003) Tyrosinase processing and intracellular trafficking is disrupted in mouse primary melanocytes carrying the underwhite (uw) mutation. A model for oculocutaneous albinism (OCA) type 4. *Journal of Cell Science* **116**, 3203–12.

Dooley C.M., Schwarz H., Mueller K.P., Mongera A., Konantz M., Neuhauss S.C., Nüsslein-Volhard C. & Geisler R. (2013) Slc45a2 and V-ATPase are regulators of melanosomal pH homeostasis in zebrafish, providing a mechanism for human pigment evolution and disease. *Pigment Cell & Melanoma Research* **26**, 205–17.

Fukamachi S., Shimada A. & Shima A. (2001) Mutations in the gene encoding B, a novel transporter protein, reduce melanin content in medaka. *Nature Genetics* **28**, 381–5.

Grønskov K., Ek J. & Brondum-Nielsen K. (2007) Oculocutaneous albinism. *Orphanet Journal of Rare Diseases* **2**, 43.

Gunnarsson U., Hellström A.R., Tixier-Boichard M., Minvielle F., Bed'hom B., Ito S., Jensen P., Rattink A., Vereijken A. & Andersson L. (2007) Mutations in *SLC45A2* cause plumage color variation in chicken and Japanese quail. *Genetics* **175**, 867–77.

Mariat D., Taourit S. & Guérin G. (2003) A mutation in the *MATP* gene causes the cream coat colour in the horse. *Genetics Selection Evolution* **35**, 119–33.

Montoliu L., Grønskov K., Wei A.H. *et al.* (2014) Increasing the complexity: new genes and new types of albinism. *Pigment Cell & Melanoma Research* **27**, 11–8.

Newton J.M., Cohen-Barak O., Hagiwara N., Gardner J.M., Davisson M.T., King R.A. & Brilliant M.H. (2001) Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4. *American Journal of Human Genetics* **69**, 981–8.

Prado-Martinez J., Hernando-Herraez I., Lorente-Galdos B. *et al.* (2013) The genome sequencing of an albino Western lowland gorilla reveals inbreeding in the wild. *BMC Genomics* **14**, 363.

Schmutz S.M. & Berryere T.G. (2007) The genetics of cream coat color in dogs. *Journal of Heredity* **98**, 544–8.

Wijesena H.R. & Schmutz S.M. (2015) A missense mutation in *SLC45A2* is associated with albinism in several small long haired dog breeds. *Journal of Heredity* **106**, 285–8.

Winkler P.A., Gornik K.R., Ramsey D.T., Dubielzig R.R., Venta P.J., Petersen-Jones S.M. & Bartoe J.T.. (2014) A partial gene deletion of *SLC45A2* causes oculocutaneous albinism in Doberman pinscher dogs. *PLoS ONE* **9**, e92127.

Xu X., Dong G.X., Hu X.S. *et al.* (2013) The genetic basis of white tigers. *Current Biology* **23**, 1031–5.

Supporting information

Additional supporting information may be found online in the supporting information tab for this article:

Table S1 Private homozygous protein-changing variants in the affected Bullmastiff.

Table S2 Control dogs.

OCA2 splice site variant in German Spitz dogs with oculocutaneous albinism

Journal:	PLoS One
Manuscript status:	published
Contributions:	supervision of genetic analyses, review and editing of manuscript

PLOS ONE

RESEARCH ARTICLE

OCA2 splice site variant in German Spitz dogs with oculocutaneous albinism

Madleina Caduff[1,2], Anina Bauer[1,2], Vidhya Jagannathan[1,2], Tosso Leeb[1,2]*

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **2** DermFocus, University of Bern, Bern, Switzerland

* Tosso.Leeb@vetsuisse.unibe.ch

Citation: Caduff M, Bauer A, Jagannathan V, Leeb T (2017) *OCA2* splice site variant in German Spitz dogs with oculocutaneous albinism. PLoS ONE 12 (10): e0185944. https://doi.org/10.1371/journal.pone.0185944

Editor: Claire Wade, University of Sydney Faculty of Veterinary Science, AUSTRALIA

Received: July 11, 2017

Accepted: September 21, 2017

Published: October 3, 2017

Data Availability Statement: Whole genome sequence data from 192 dogs are available from the European Nucleotide Archive (ENA). The complete list of accession numbers is given in S4 Table. The full length canine *OCA2* mRNA sequence is available under accession LT844587 from the ENA.

Funding: This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738; http://www.snf.ch) and the Albert-Heim Foundation (no. 105; http://www.albert-heim-stiftung.ch/). The funders had no role

Abstract

We investigated a German Spitz family where the mating of a black male to a white female had yielded three puppies with an unexpected light brown coat color, lightly pigmented lips and noses, and blue eyes. Combined linkage and homozygosity analysis based on a fully penetrant monogenic autosomal recessive mode of inheritance identified a critical interval of 15 Mb on chromosome 3. We obtained whole genome sequence data from one affected dog, three wolves, and 188 control dogs. Filtering for private variants revealed a single variant with predicted high impact in the critical interval in *LOC100855460* (XM_005618224.1: c.377+2T>G LT844587.1:c.-45+2T>G). The variant perfectly co-segregated with the phenotype in the family. We genotyped 181 control dogs with normal pigmentation from diverse breeds including 22 unrelated German Spitz dogs, which were all homozygous wildtype. Comparative sequence analyses revealed that *LOC100855460* actually represents the 5'-end of the canine *OCA2* gene. The CanFam 3.1 reference genome assembly is incorrect and separates the first two exons from the remaining exons of the *OCA2* gene. We amplified a canine *OCA2* cDNA fragment by RT-PCR and determined the correct full-length mRNA sequence (LT844587.1). Variants in the *OCA2* gene cause oculocutaneous albinism type 2 (OCA2) in humans, pink-eyed dilution in mice, and similar phenotypes in corn snakes, medaka and Mexican cave tetra fish. We therefore conclude that the observed oculocutaneous albinism in German Spitz is most likely caused by the identified variant in the 5'-splice site of the first intron of the canine *OCA2* gene.

Introduction

Oculocutaneous albinism (OCA) summarizes a group of inherited disorders of melanin synthesis, characterized by hypopigmentation in skin, hair and eyes [1]. Six of the seven types of human OCA are assigned to variants in a specific gene: OCA1 (*TYR*), OCA2 (*OCA2*), OCA3 (*TYRP1*), OCA4 (*SLC45A2*), OCA6 (*SLC24A5*) and OCA7 (*LRMDA*) [1,2]. Variants in *OCA2* cause OCA2 in humans (OMIM #203200) and pink-eyed dilution in the mouse [3]. Up to 150 human and 100 murine *OCA2* variants have been reported [4,5]. Additionally, amelanism in corn snakes and albinism in medaka and Mexican cave tetra fish are also caused by variants in *OCA2* [6–8].

in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

The phenotype in human OCA2 patients is variable. The amount of residual cutaneous pigment varies and typically increases with age. Iris color is also variable. As with other types of oculocutaneous albinism, hypopigmentation of the iris leading to iris translucency, reduced pigmentation of the retinal pigment epithelium, foveal hypoplasia, reduced visual acuity, refractive errors, and sometimes a degree of color vision impairment may occur. The human OCA2 phenotype is the most common type of OCA in Africans and has also been termed brown OCA in Africans [1,2].

The human *OCA2* gene spans 380 kb on chromosome 15 and the major transcript isoform contains 24 exons [9]. The OCA2 melanosomal transmembrane protein, formerly called P-protein, is a putative anion transporter with 12 transmembrane domains. OCA2 plays a role in melanosome biogenesis, melanosomal pH regulation, and eumelanin synthesis [3,5,10,11]. It is required for the normal processing and transport of other melanosomal proteins, such as TYR and TYRP1 [1].

The aim of the current study was to unravel the molecular genetics for an oculocutaneous albinism phenotype in a family of German Spitz dogs.

Results

Phenotype characterization

A German Spitz family of the Giant Spitz variety was brought to our attention where the mating of a black sire and a white bitch had produced a litter with six puppies. Three puppies had the expected black coat color, while the other three puppies, two males and one female, were of a light brown color, which could not be explained by the genotypes of the known basic coat color loci (Fig 1). Father, mother and the unaffected littermates had dark-colored noses and brown to amber eyes, whereas the three brown puppies had light lips and noses and blue eyes that turned into green with age. Dilated pupils appeared of reddish color. The coat color darkened somewhat with age. The owner reported that the affected puppies used to squint in bright sunlight (photophobia) and had difficulties to perceive hand signals in bright sunlight. Breeders' reports indicated that dogs with similar phenotypes had previously been noticed in this line of dogs. These data suggested a monogenic autosomal recessive mode of inheritance.

Mapping of the causative locus

We genotyped all members of the family on the Illumina canineHD SNP chip and employed a combined linkage and homozygosity approach in the family with its six offspring. Parametric linkage analysis was performed assuming a fully penetrant, monogenic autosomal recessive inheritance of the trait. This analysis identified 10 segments $\geq$ 1 Mb that showed positive LOD scores (S1 Table). Autozygosity mapping in the three affected dogs yielded three homozygous regions with shared alleles (S2 Table). Only one genome segment, Chr3:28,747,944–43,731,542, showed both linkage and homozygosity (Fig 2).

Identification of the causative variant

We resequenced the whole genome of one affected dog and called single nucleotide as well as indel variants with respect to the reference genome of a non-affected Boxer (CanFam 3.1). The genotypes of the affected German Spitz were further compared with 188 dog genomes from various breeds and three wolf genomes that had been sequenced in the course of other studies. We hypothesized that the causative variant should be completely absent from all other genomes except the German Spitz with oculocutaneous albinism.

Fig 1. Coat color phenotypes in the investigated Giant Spitz family. (A) White mother and black father of the litter. **(B)** Pictures of two affected puppies and one unaffected sibling on the left at one week of age. **(C, E)** Affected dog GS103 with green eyes and a light nose at 4 and 5.5 months of age, respectively. **(D, F)** Affected dog GS104 at 4 and 7 months of age, respectively.

https://doi.org/10.1371/journal.pone.0185944.g001

Within the critical interval, this filtering process yielded only one single private variant with predicted impact on a protein-coding gene (Table 1). This variant affected the 5'-splice site of the single intron of *LOC100855460* (Chr3:31,715,704A> C; XM_005618224.1:c.377+2T> G).

We confirmed the splice site variant by Sanger sequencing and genotyped the 8 family members, 22 unrelated German Spitz dogs (8 Giant Spitz dogs, 7 Miniature Spitz dogs and 7 Pomeranians), and 159 control dogs from various other breeds (Fig 3A). The genotypes at this variant showed perfect co-segregation with the phenotype (Fig 3B). None of the 181 normally pigmented control dogs carried the mutant allele (S3 Table).

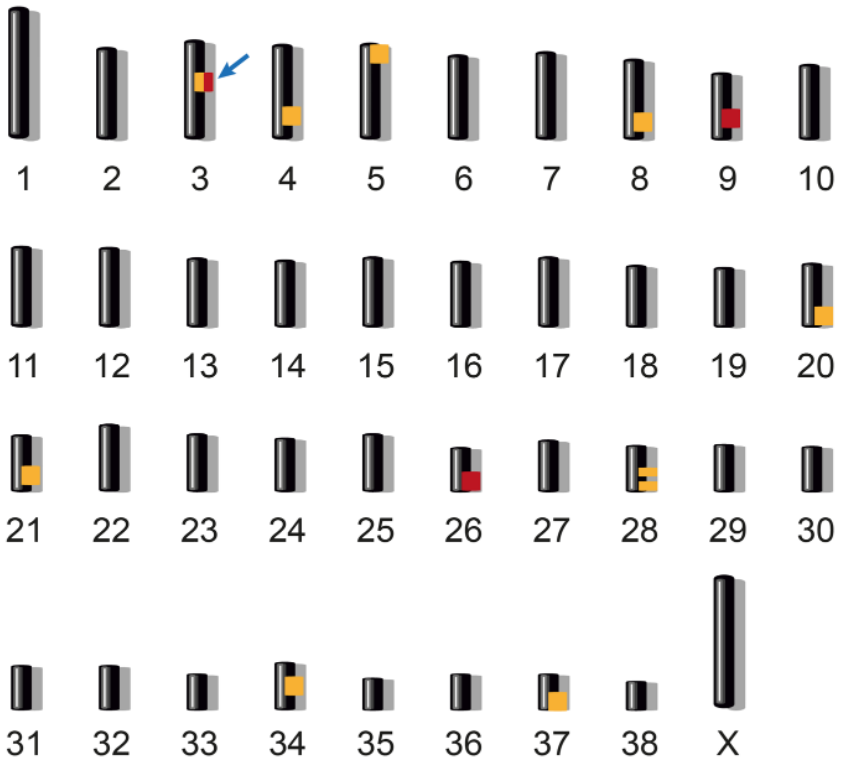


Fig 2. Combined linkage and homozygosity mapping. We performed parametric linkage analysis for a recessive trait in all eight family members and homozygosity mapping across the three affected German Spitz dogs. Ten linked genome segments are indicated in orange and three homozygous segments with shared alleles are indicated in red. Only one region on chromosome 3 showed both linkage and homozygosity and was considered the critical interval (arrow). Specifically, this ~15 Mb region corresponded to Chr3:28,747,944–43,731,542 (CanFam 3.1 assembly).

https://doi.org/10.1371/journal.pone.0185944.g002

LOC100855460 represents the 5'-end of the *OCA2* gene

As the *LOC100855460* splice site variant represented our best candidate causative variant for the observed oculocutaneous albinism phenotype, we performed BLAST searches with *LOC100855460* to identify putative homologs in other mammalian species. These analyses revealed that *LOC100855460* corresponds to the first two exons of the human *OCA2* gene. The

Table 1. Variants detected by whole genome resequencing of an affected German Spitz.

Filtering step	Number of variants
Variants in the whole genome[a]	540,127
Variants in the critical 15 Mb interval on chromosome 3	5,855
Variants in the critical interval that were absent from 191 other dog genomes	199
Protein-changing variants in the whole genome[a]	2,317
Protein-changing variants in the 15 Mb critical interval on chromosome 3	15
Protein-changing variants in the critical interval, absent from 191 other dog genomes	1

[a]The sequences were compared to the reference genome (CanFam 3.1) from a Boxer. Only variants that passed GATK quality filters and were homozygous in the affected German Spitz GS104 are reported. Protein-changing variants were classified based on the NCBI annotation release 104.

https://doi.org/10.1371/journal.pone.0185944.t001

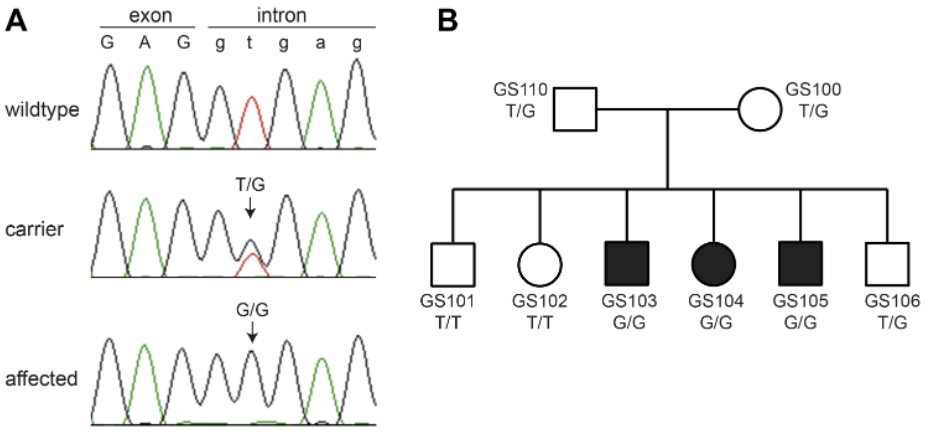


Fig 3. Genotypes of the German Spitz litter. (A) Sanger sequencing chromatograms illustrate sequences of a homozygous wildtype dog, a heterozygous carrier and a homozygous affected dog. The variant replaced the essential thymine at position +2 of the 5'-splice site with a guanine. **(B)** Pedigree of the Giant Spitz litter. Genotypes at the splice site variant show the expected co-segregation with the phenotype.

https://doi.org/10.1371/journal.pone.0185944.g003

currently incorrect canine gene annotation in the region is most likely due to an assembly error in the CanFam3.1 reference genome that contains a ~300 kb segment of genomic DNA in inverted orientation (Fig 4).

Experimental confirmation of the proposed canine *OCA2* gene structure

We confirmed the suspected genome assembly error by long-range PCR on genomic DNA. We then designed an RT-PCR primer pair with a forward primer in *LOC100855460* and a reverse primer in the canine *OCA2* gene. RT-PCR and subsequent Sanger sequencing of the product demonstrated that *LOC100855460* and *OCA2* are indeed parts of the same transcript.

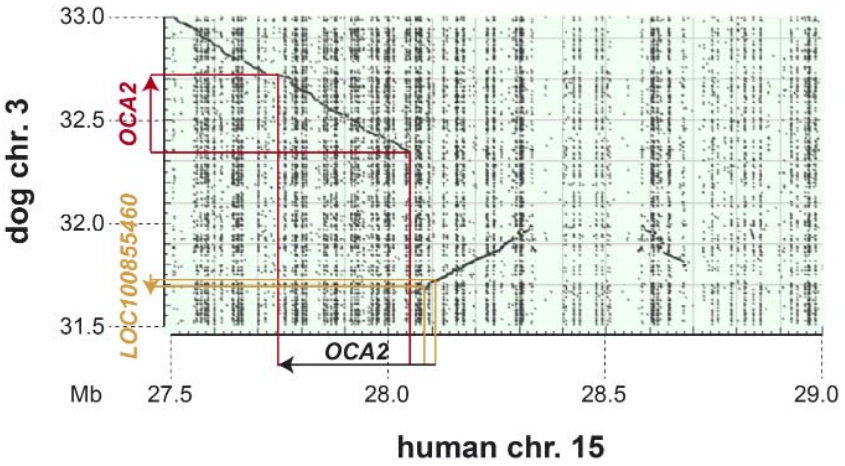


Fig 4. Dot plot of human Chr15:27,500,001–29,000,000 against dog Chr3:31,500,001–33,000,000. The human *OCA2* gene spans 380 kb. The dot plot illustrates that ~300 kb of dog reference sequence are inverted with respect to the human sequence due to an assembly error in the current CanFam 3.1 assembly. Due to this assembly error the first two exons of the canine *OCA2* gene are currently in the wrong orientation and annotated as *LOC100855460*.

https://doi.org/10.1371/journal.pone.0185944.g004

Based on the RT-PCR results and public RNA-seq data, we determined a full-length canine *OCA2* mRNA sequence and submitted it to the European Nucleotide Archive (accession number LT844587.1). This transcript contains an open reading frame of 2,532 nucleotides encoding a protein of 844 amino acids that shows 83% identity to the 838 amino acid human OCA2 melanosomal transmembrane protein, isoform 1 (NP_000266.2). Most of the sequence differences are located in the first 180 residues at the N-terminus.

Based on our revised canine *OCA2* full-length sequence, the candidate causative variant for oculocutaneous albinism in dogs should be designated *OCA2*:LT844587.1:c.-45+2T> G. It affects the 5'-splice site of the first intron of the canine *OCA2* gene.

Discussion

In this study, we identified a splice site variant in the canine *OCA2* gene as most likely cause for an oculocutaneous albinism phenotype in dogs. There are two main arguments supporting the claimed causality: The variant affects the conserved GT dinucleotide at the 5'-splice site of an AG-GT intron [12]. Such a genomic alteration is not compatible with the normal splicing process and is predicted to lead to a loss of function of the mutant gene. Secondly, we provided evidence that the splice site variant affects the first intron of the canine *OCA2* gene. The functional role of this gene in pigmentation is well known and numerous genetic variants in humans, mice, and other vertebrates, lead to oculocutaneous albinism phenotypes that are very similar to the phenotypes observed in the dogs in this study [3–8]; (OMIA 000202–7994, OMIA 000202–94885). *OCA2* mutant alleles were reported to mainly block eumelanin synthesis with a less pronounced effect on pheomelanin synthesis [4]. Humans with OCA2 show a dramatic reduction of eumelanin in skin, eye and hair, resulting in cream-colored skin and yellow-straw hair, which darkens with age [13]. Melanosomes with an irregular shape were observed in p-deficient mice [14], and measurements of melanin content revealed a great decrease of eumelanin, but not pheomelanin levels [15,16]. *OCA2* expression was only found in the black dorsal, but not in the yellow ventral part of black-and-tan (a^t/a^t) mice, whereas it was present in the ventral black skin of non-agouti (a/a) mice [3]. These findings are in line with the phenotype of the *OCA2* deficient dogs in the present study. Based on their genotypes at the other coat color loci, the three affected dogs would have been expected to be black due to the presence of the dominant K^B allele (*E/e*; A^w/a; K^B/k^y; *B/B*, *D/-*). It is not clear whether the residual light brown pigment in the affected dogs represents a mixture of eumelanin and pheomelanin or whether this is an abnormal melanin with a brownish color.

Accumulation of pigment with age has been described in human OCA2 patients [17] and was observed in the affected dogs as the eye color changed from blue into light green and the coat color darkened. The observed photophobia and translucent irises of the affected dogs agrees with the description of visual abnormalities in human OCA patients [1,18]. Due to the photophobia, such dogs should not be bred on purpose and carrier x carrier matings should be avoided in the future.

This study highlights the current status of genomic resources in dogs and many other domestic animal species. Although whole genome sequencing has been shown to represent a powerful technology when it comes to identifying causative variants for Mendelian traits, the bioinformatic data analysis is hindered by imperfect reference genome assemblies and gene annotations. With the current resources, a significant amount of human expert data interpretation is required. In the future, with the availability of better reference genomes, it may be expected that similar analyses will become easier and involve a greater extent of automated data analysis.

In conclusion, we identified a variant in the conserved 5'-splice site of the first intron of the canine *OCA2* gene as most likely cause for the observed oculocutaneous albinism in a Giant Spitz family. This will facilitate genetic testing to avoid the accidental breeding of further puppies with this phenotype. Our work has revealed an assembly error in the CanFam3.1 assembly and we provide a revised full-length sequence of the canine *OCA2* cDNA.

Materials and methods

Ethics statement

All animal experiments were performed according to the local regulations. The dogs in this study were examined with the consent of their owners. The study was approved by the "Cantonal Committee For Animal Experiments" (Canton of Bern; permits 75/16 and 38/17).

DNA samples and genotyping

We obtained EDTA-blood samples from both parents and all 6 puppies of the German Spitz litter (Giant Spitz variety). We isolated genomic DNA with the Maxwell RSC Whole Blood DNA Kit and the Maxwell RSC Instrument (Promega). Genotyping of the two parents and six littermates was performed by GeneSeek/Neogen on the Illumina CanineHD BeadChip containing 220,853 markers. We also used 181 canine control DNA samples from the Vetsuisse Biobank that had been collected during the course of other projects.

Linkage analysis and homozygosity mapping

The genotype data of the 8 family members was used for a parametric linkage analysis. The call rate was > 95% for all dogs. Using PLINK v 1.07 [19], markers that were non-informative, located on the sex chromosomes, or missing in any of the 8 dogs, had Mendel errors, or a minor allele frequency < 0.3125, were removed. The final pruned dataset contained 34,356 markers. An autosomal recessive inheritance model with full penetrance, a disease allele frequency of 0.5 and the Merlin software [20] were applied to test for linkage. Intervals with positive LOD scores and $\alpha = 1$ were retained for further analysis.

For homozygosity mapping, the genotype data of the three affected dogs were used. Markers that were missing in one of the three cases, markers on the sex chromosomes and markers with Mendel errors in the family were excluded. Using the—homozyg and—homozyg-group options in PLINK, extended regions of homozygosity > 1 Mb were identified. The homozygous intervals in the three cases were intersected with the linked intervals to define minimal critical intervals.

Whole-genome sequencing

An Illumina TruSeq PCR-free library with an insert size of 350 bp was prepared from one affected dog (GS104). The library was sequenced at 32x genome coverage using 2 x 150 bp reads on an Illumina HiSeq 3000 instrument. Single nucleotide and small indel variants with respect to the CanFam3.1 canine reference genome assembly were called as described [21]. Variants private to GS104 were identified by filtering variants that were contained in 3 wolf and 188 control dog genomes sequenced for previous projects (S4 Table). Private variants were prioritized according to their predicted impact using SNPeff [22] and the NCBI annotation release 104.

Sanger sequencing

We used Sanger sequencing to confirm the Illumina sequencing results and to perform targeted genotyping for the splice site variant in *LOC100855460*. AmpliTaqGold360Mastermix (Applied Biosystems) and the primer pair GTCTGGCCTTTCCGTGAG (forward primer) and CGAAGCTTGTGCTCAATGTC (reverse primer) were used to amplify the region by PCR. PCR products were directly sequenced on an ABI 3730 capillary sequencer (Applied Biosystems) after treatment with exonuclease I and shrimp alkaline phosphatase. The reverse primer was used as sequencing primer. We analyzed the sequence data with Sequencher 5.1 (GeneCodes).

Long-range PCR

A long-range PCR to was performed using the primer pair GGCAAACTTGGGAGTGGTAA (Chr3:31,659,781–31,659,762) and CCCCCTCAAATAAACCATGA (Chr3:32,346,375–32,346,356) and the SequalPrep™ Long PCR kit (Invitrogen). The PCR fragments were analyzed using FragmentAnalyzer™ (Advanced Analytical).

RT-PCR

We isolated total RNA from a skin biopsy of a healthy control dog using QIAzol and RNeasy spin columns according to the manufacturer's recommendations (Qiagen). RNA samples were treated with RNase-free DNase to remove contaminations with genomic DNA. Reverse transcription was carried out using an oligo-dT primer, and Superscript[1] IV Reverse Transcriptase according to the manufacturer's recommendations (Invitrogen). PCR and sequencing were performed with 2 μl of the synthesized cDNA and the primer pair TTCTTTCTGGCTGACCTCGT (forward primer, *OCA2* exon 2, Chr3:31,681,124–31,681,105) and ATGCACCATGACCCTTTC TC (reverse primer, *OCA2* exons 4 & 3, Chr3:32,366,619–32,366,608 & Chr3:32,362,337–32,362,330), using the forward primer as a sequencing primer. For determination of the 5'- and 3'-ends of the *OCA2* mRNA sequence, we utilized publicly available RNA-seq data from canine nasal skin (ENA project accession PRJEB14109, sample accession SAMEA4412813, Lab ID LA1666).

Supporting information

S1 Table. Results of linkage analysis.
(XLSX)

S2 Table. Results of homozygosity mapping.
(XLSX)

S3 Table. Genotypes of control dogs.
(XLSX)

S4 Table. Accession numbers of 192 dog/wolf genome sequences.
(XLSX)

Acknowledgments

The authors would like to thank the breeder Barbara Tuschl and the current owners of the dogs for donating samples and pictures and for sharing pedigree information of their dogs. The authors also wish to thank Nathalie Besuchet, Muriel Fragnière, and Sabrina Schenk for expert technical assistance. We acknowledge collaborators of the Dog Biomedical Variant Database Consortium (DBVDC), Gus Aguirre, Catherine André, Danika Bannasch, Doreen

Becker, Cord Drögemüller, Kari Ekenstedt, Oliver Forman, Steve Friedenberg, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Vidhya Jagannathan, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Leonardo Murgiano, Anita Oberbauer, Sheila Schmutz, Jeffrey Schoenebeck, Kim Summers, Frank van Steenbeck, Claire Wade for sharing dog genome sequence data from control dogs and wolves. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure.

Author Contributions

Data curation: Vidhya Jagannathan.

Funding acquisition: Tosso Leeb.

Investigation: Madleina Caduff, Anina Bauer, Vidhya Jagannathan.

Methodology: Anina Bauer, Vidhya Jagannathan.

Supervision: Anina Bauer, Tosso Leeb.

Writing – original draft: Madleina Caduff.

Writing – review & editing: Tosso Leeb.

References

1. Grønskov K, Ek J, Brondum-Nielsen K. Oculocutaneous albinism. Orphanet J Rare Dis. 2007; 2: 43. https://doi.org/10.1186/1750-1172-2-43 PMID: 17980020
2. Montoliu L, Grønskov K, Wei AH, Martínez-García M, Fernández A, Arveiler B, et al. Increasing the complexity: new genes and new types of albinism. Pigment Cell Melanoma Res. 2014; 27: 11–18. https://doi.org/10.1111/pcmr.12167 PMID: 24066960
3. Rinchik EM, Bultman SJ, Horsthemke B, Lee ST, Strunk KM, Spritz RA et al. A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. Nature. 1993; 361: 72–76. https://doi.org/10.1038/361072a0 PMID: 8421497
4. Mártinez-García M, Montoliu L. Albinism in Europe. *J Dermatol*. 2013; 40: 319–324. https://doi.org/10.1111/1346-8138.12170 PMID: 23668539
5. Brilliant M. The mouse p (pink-eyed dilution) and human P genes, oculocutaneous albinism type 2 (OCA2), and melanosomal pH. Pigment Cell Res. 2001; 14: 86–93. PMID: 11310796
6. Saenko SV, Lamichhaney S, Barrio AM, Rafati N, Andersson L, Milinkovitch MC. Amelanism in the corn snake is associated with the insertion of an LTR-retrotransposon in the *OCA2* gene. Sci Rep. 2015; 5: 17118. https://doi.org/10.1038/srep17118 PMID: 26597053
7. Fukamachi S, Asakawa S, Wakamatsu Y, Shimizu N, Mitani H, Shima A. Conserved function of medaka pink-eyed dilution in melanin synthesis and its divergent transcriptional regulation in gonads among vertebrates. Genetics. 2004; 168: 1519–1527. https://doi.org/10.1534/genetics.104.030494 PMID: 15579703
8. Protas ME, Hersey C, Kochanek D, Zhou Y, Wilkens H, Jeffery WR, et al. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nature Genet. 2006; 38: 107–111. https://doi.org/10.1038/ng1700 PMID: 16341223
9. NCBI Gene website describing the features of the human OCA2 gene (annotation release 108). https://www.ncbi.nlm.nih.gov/gene/4948.
10. Rosemblat S, Durham-Pierre D, Gardner JM, Nakatsu Y, Brilliant MH, Orlow SJ. Identification of a melanosomal membrane protein encoded by the pink-eyed dilution (type II oculocutaneous albinism) gene. Proc Natl Acad Sci USA. 1994; 91: 12071–12075. PMID: 7991586
11. Hirobe T. How are proliferation and differentiation of melanocytes regulated?. Pigment Cell Melanoma Res. 2011; 24: 462–478. https://doi.org/10.1111/j.1755-148X.2011.00845.x PMID: 21375698
12. Burset M, Seledtsov IA, Solovyev VV. SpliceDB: database of canonical and non-canonical mammalian splice sites. Nucleic Acids Res. 2001; 29: 255–259. PMID: 11125105

13. Manga P, Orlow SJ. The Pink-Eyed Dilution Gene and the Molecular Pathogenesis of Tyrosinase-Positive Albinism (OCA2). J Dermatol. 1999; 26: 738–747. PMID: 10635616
14. Russell ES. A quantitative histological study of the pigment found in the coat-color mutants of the house mouse. IV. The nature of the effects of genic substitution in five major allelic series. *Genetics*. 1949; 34: 146–166.
15. Ozeki H, Ito S, Wakamatsu K, Hirobe T. Chemical characterization of hair melanins in various coat-color mutants of mice. J Invest Dermatol. 1995; 105: 361–366. PMID: 7665913
16. Prota G, Lamoreux M, Muller J, Kobayashi T, Napolitano A, Vincensi MR, et al. Comparative analysis of melanins and melanosomes produced by various coat color mutants. *Pigment Cell Res*. 1995; 8: 153–163. PMID: 7567792
17. King RA, Creel D, Arvenka J, Okord AN, Witkop CJ. Albinism in Nigeria with delineation of new recessive oculocutaneous type. Clin Genet. 1980; 17: 259–270. PMID: 6768477
18. Lee ST, Nicholls RD, Bundey S, Laxova R, Musarella M, Spritz RA. Mutations of the P gene in oculocutaneous albinism, ocular albinism, and Prader-Willi syndrome plus albinism. N Engl J Med. 1994; 330: 529–534. https://doi.org/10.1056/NEJM199402243300803 PMID: 8302318
19. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81: 559–575. https://doi.org/10.1086/519795 PMID: 17701901
20. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. 2002; 30: 97–101. https://doi.org/10.1038/ng786 PMID: 11731797
21. Bauer A, Waluk DP, Galichet A, Timm K, Jagannathan V, Sayar BS, et al. A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. PLoS Genet. 2017; 13: e1006651. https://doi.org/10.1371/journal.pgen.1006651 PMID: 28249031
22. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012; 6: 80–92.

A nonsense variant in the *ST14* gene in Akhal-Teke horses with naked foal syndrome

Journal:	G3: Genes, Genomes, Genetics
Manuscript status:	published
Contributions:	genetic analyses, original draft, review and editing of manuscript, illustrations

G3 Genes | Genomes | Genetics

INVESTIGATION

A Nonsense Variant in the *ST14* Gene in Akhal-Teke Horses with Naked Foal Syndrome

Anina Bauer,[*,†,‡] Theresa Hiemesch,[*,§] Vidhya Jagannathan,[*,†,‡] Markus Neuditschko,[‡,] Iris Bachmann,[‡,**] Stefan Rieder,[‡,**] Sofia Mikko,[††] M. Cecilia Penedo,[‡‡] Nadja Tarasova,[§§,***] Martina Vitková,[†††] Nicolò Sirtori,[‡‡‡] Paola Roccabianca,[§§§] Tosso Leeb,[*,†,‡,1] and Monika M. Welle[†,****]**

*Institute of Genetics, and ****Institute of Animal Pathology, Vetsuisse Faculty, and [†]DermFocus, University of Bern, 3001, Switzerland, [‡]Swiss Competence Center of Animal Breeding and Genetics, University of Bern, Bern University of Applied Sciences HAFL & Agroscope, 3001, Switzerland, [§]Institute of Animal Breeding and Genetics, University of Göttingen, 37075, Germany, **Agroscope, Swiss National Stud Farm, 1580 Avenches, Switzerland, [††]Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden, [‡‡]Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, California 95616, [§§]Russian Akhal-Teke Association, and ***International Akhal-Teke Association, 115470 Moscow, Russia, [†††]Equine Veterinary Practice, 91601 Stará Turá, Slovakia, [‡‡‡]Equine Veterinary Practice, 29010 Agazzano, Italy, and [§§§]Department of Veterinary Medicine, University of Milan, 20133, Italy

ORCID IDs: 0000-0002-6765-4012 (A.B.); 0000-0003-0553-4880 (T.L.)

ABSTRACT Naked foal syndrome (NFS) is a genodermatosis in the Akhal-Teke horse breed. We provide the first scientific description of this phenotype. Affected horses have almost no hair and show a mild ichthyosis. So far, all known NFS affected horses died between a few weeks and 3 yr of age. It is not clear whether a specific pathology caused the premature deaths. NFS is inherited as a monogenic autosomal recessive trait. We mapped the disease causing genetic variant to two segments on chromosomes 7 and 27 in the equine genome. Whole genome sequencing of two affected horses, two obligate carriers, and 75 control horses from other breeds revealed a single nonsynonymous genetic variant on the chromosome 7 segment that was perfectly associated with NFS. The affected horses were homozygous for *ST14*:c.388G>T, a nonsense variant that truncates >80% of the open reading frame of the *ST14* gene (p.Glu130*). The variant leads to partial nonsense-mediated decay of the mutant transcript. Genetic variants in the *ST14* gene are responsible for autosomal recessive congenital ichthyosis 11 in humans. Thus, the identified equine *ST14*:c.388G>T variant is an excellent candidate causative variant for NFS, and the affected horses represent a large animal model for a known human genodermatosis. Our findings will enable genetic testing to avoid the nonintentional breeding of NFS-affected foals.

KEYWORDS
Equus caballus
dermatology
skin
hair
genodermatosis
whole genome sequencing

Spontaneous mutants in domestic animals represent valuable animal models for genetic research. Several such mutants, with either missing or altered hair characteristics, have contributed to our knowledge on hair follicle development and hair growth. These include hairless dog breeds, such as the Chinese Crested, Mexican Hairless, and Peruvian Hairless dogs, which show an ectodermal dysplasia phenotype characterized by missing hair and altered dentition due to a genetic variant in the *FOXI3* gene (Drögemüller *et al.* 2008). A genetic variant in the *KRT71* gene is responsible for the hairless Sphynx cats, and other variants in *KRT71* cause curly hair in dogs and cats (Cadieu *et al.* 2009; Gandolfi *et al.* 2010, 2013). In contrast to these breed-defining characters, several hair-related traits with very severe effects on other organ systems are actively selected against. An example for this latter group is the "hairlessness with short life expectancy" in *FOXN1* mutant cats that involves a severe immunodeficiency, in addition to the missing hair, and thus resembles *Foxn1* mutant *nude* mice and rats (Nehls *et al.* 1994; Abitbol *et al.* 2015).

doi: https://doi.org/10.1534/g3.117.039511
Manuscript received January 13, 2017; accepted for publication February 18, 2017; published Early Online February 22, 2017.

Supplemental material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.117.039511/-/DC1.
[1]Corresponding author: Institute of Genetics, University of Bern, Bremgartenstrasse 109a, 3001 Bern, Switzerland. E-mail: Tosso.Leeb@vetsuisse.unibe.ch

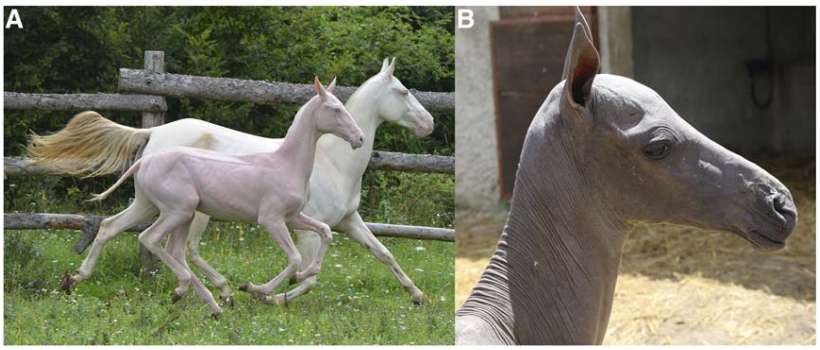


Figure 1 NFS phenotype. (A) An NFS affected colt (case 1) in front of its nonaffected mother. The affected colt has almost no hair. Both horses are of cremello coat color (*SLC45A2*:c.457G>A), which leads to a very strong dilution of the pigmentation but is unrelated to the NFS phenotype. (B) NFS affected filly (case 2).

The Akhal-Teke horse breed has its origin in central Asia and is known for high endurance and the characteristic metallic shine of its hair (Leisson *et al.* 2011). The Akhal-Teke studbook is closed since 1932. Due to the limited gene pool and the practice of line breeding, there is a high degree of inbreeding in the Akhal-Teke breed, which favors the appearance of recessive genetic disorders (Leroy and Baumung 2011). The naked foal syndrome (NFS) is a genetic disorder in the Akhal-Teke breed that is known among breeders (Kuznetcova *et al.* 2006; Kuznetcova 2006). However, according to our knowledge, NFS has never been described in the scientific literature. Horses with this disorder are born hairless, and often die within days to months after birth. However, the reason for these early deaths is not known, and some hairless foals have survived up to 2.5 yr. The first records of hairless Akhal-Teke foals date back to 1938, and since then the number of such foals has increased steadily. Many horses with NFS might have been registered as stillborn or weak born, or not been reported at all (Kuznetcova *et al.* 2006).

The aim of the present study was to identify the causative genetic variant for NFS, and to develop a genetic test enabling the identification of carriers. As this phenotype has never been fully described in the scientific literature, we also present a preliminary qualitative characterization of the NFS phenotype.

MATERIALS AND METHODS

Ethics statement

Animal work consisted of collecting blood samples and skin biopsies of privately owned horses with owners' consent, in accordance with the relevant local guidelines (permits no. BE31/13 and BE75/16).

Samples and genotyping

The study included 206 Akhal-Teke horses, including five NFS affected horses and 10 obligate carriers (= owner-reported parents of NFS affected animals, Supplemental Material, Table S1). Genomic DNA was isolated from EDTA blood, or hair samples, or, in the case of one deceased affected foal, from a hoof that had been archived. In addition to the Akhal-Teke horses, we used genomic DNA samples from 400 control horses of different breeds that had been collected in the course of other projects and were stored in our biobank. We additionally collected blood from a carrier in a PAXgene blood RNA tube to enable isolation of RNA. Genotyping of four affected horses and four obligate carriers was performed by GeneSeek/NeoGene on the Affymetrix equine 670 k SNP array containing 670,796 evenly distributed markers.

Skin biopsy sampling and histopathology

Three 8-mm punch biopsies of an NFS affected, an obligate carrier, and a control Akhal-Teke horse were collected and fixed in 4% buffered formalin. Wedge biopsies were taken from a second NFS affected Akhal-Teke foal after euthanasia. We processed the tissue, and stained the skin sections with hematoxylin and eosin (H&E) prior to histopathological examination by a board certified veterinary pathologist (M.W.).

Linkage analysis

We used genotype data from an Akhal-Teke family consisting of three NFS affected half-siblings, and their four parents, to perform a parametric linkage analysis (Figure S1). For all horses, the call rate was >95%. Using PLINK v1.07 (Purcell *et al.* 2007), we removed markers that were noninformative, located on the sex chromosomes, missing in any of the seven horses, had Mendel errors, or a minor allele frequency <0.2. The final pruned dataset contained 210,556 markers. To analyze the data for parametric linkage, we applied an autosomal recessive inheritance model with full penetrance and a disease allele frequency of 0.7, and the Merlin software (Abecasis *et al.* 2002).

Homozygosity mapping

Affymetrix equine 670 k SNP genotypes were available for four affected horses. One of these cases was excluded from further analysis because of a low call rate of <95%. We further excluded markers that were missing in one of the three remaining cases, markers on the sex chromosomes, and markers with Mendel errors in the family (see above). Using the --homozyg and --homozyg-group options in PLINK, we searched for extended regions of homozygosity >1 Mb. We visually inspected the raw genotypes to compile the exact boundaries of the homozygous intervals in an Excel file.

Whole genome sequencing of two affected horses and two carriers

We prepared illumina PCR-free TruSeq fragment libraries with an insert size of ~350 bp from two cases (AKT001 and AKT006) and two carriers (AKT003 and AKT004) and collected 19× to 33× coverage 2 × 150 bp paired-end data on an Illumina HiSeq 3000 instrument. Read mapping and variant calling was done as described previously (Murgiano *et al.* 2016).

Sanger sequencing

We used Sanger sequencing to genotype all available Akhal-Teke horses and 400 control horses of different breeds for the *ST14*:c.388G>T variant. A genomic fragment containing the variant position was PCR-amplified from genomic DNA with the primers F, CTGAGAG CAGAGGGTCAAGG, and R, GTGCACTGGCTCTGTGACTG, using AmpliTaqGold360Mastermix (Life Technologies). We directly sequenced the PCR products using the PCR primers on an ABI 3730 capillary sequencer (Life Technologies) after treatment with

exonuclease I and shrimp alkaline phosphatase. Sequence data were analyzed using Sequencher 5.1 (GeneCodes).

RNA isolation and RT-PCR

RNA was isolated from the PAXgene blood of a carrier using the PAXgene Blood RNA kit (PreAnalytiX, Qiagen), and a reverse transcription PCR was performed using Superscript IV reverse transcriptase (Life Technologies). Subsequently, a cDNA fragment was PCR amplified using the primers F, CTGGCCAATAAGGTGAAGGA and R, CCTGTGTGGTCTGTGCTGTT with AmpliTaqGold360Mastermix (Life Technologies). The RT-PCR product and the genomic PCR product described above were sequenced using the same primer SEQ, AAAGCCACCACCGAGGTC, for the sequencing reaction.

Reference genome assembly and gene annotation

The horse EquCab 2.0 reference genome assembly was used for all analyses. Numbering within the equine *ST14* gene corresponds to the NCBI accessions XM_005611718.2 (mRNA) and XP_005611775.1 (protein). Numbering within the human gene and protein corresponds to the accessions NM_021978.3 and NP_068813.1, respectively.

Data availability

Figure S1 illustrates the pedigrees of the NFS-affected horses. Table S1 gives additional information on the five NFS-affected and 10 obligate carrier horses used in this study. Genotypes for the *ST14*:c.388G>T variant from 154 Akhal-Teke horses and 400 control horses are listed in Table S2. Table S3 lists genome regions $ 1 Mb that showed positive LOD scores in the linkage analysis. In Table S4, homozygous genome regions $ 1 Mb with shared alleles among three NFS affected Akhal-Teke horses are given. The genome sequencing data of the Akhal-Teke horses were deposited in the European Nucleotide Archive, under accession PRJEB14779. Accession numbers of sequencing data from control horses can be found in Table S5.

RESULTS

Clinical and pathological phenotype

Case 1: We examined two Akhal-Teke horses with NFS. Case 1 was a male cremello foal born in March 2014 (Figure 1A). This horse was still alive at the time of manuscript revision (January 2017, 2 yr 10 months). Compared to nonaffected horses of the same age, and raised at the same stud, it had a growth delay and was small for its age. The horse was alopecic, with only sparse and thin body hairs. In both fore and hind limbs, the proximal parts were completely alopecic, while the hair density increased toward the distal ends of the limbs. Mane and hairs at the tail were sparse or absent. Whiskers were present but sparse, curly, and abnormally short. Eyelashes were missing. The skin was dry and scaly (xerosis cutis) in some body parts. Furthermore, a persisting increase in tear flow (epiphora) was reported by the owner. Multifocal scars and erosive lesions were present, possibly due to the missing physical protection of a normal hair coat. We did not observe any abnormalities in the teeth or hooves.

Case 2: Case 2 was a female, born in June 2016 (Figure 1B). The skin and hair phenotype closely resembled that of case 1 (Figure 2). Hooves and teeth were also normal. This foal had to be killed at 21 d of age due to a spontaneous leg fracture. A full necropsy on the killed foal was performed. Gross findings revealed a mild internal hydrocephalus, a heart defect (dysplasia of the tricuspidal valve), and severely altered lymphoid organs. The histological findings in the lymphoid organs were consistent with a defect in primary immune organ development, and the specific immune response. Specifically, the thymus lacked cortico-medullary organization, and was characterized by absence of Hassal corpuscles. Abnormal or absent T cell zones were noted in the spleen and the peripheral lymph nodes.

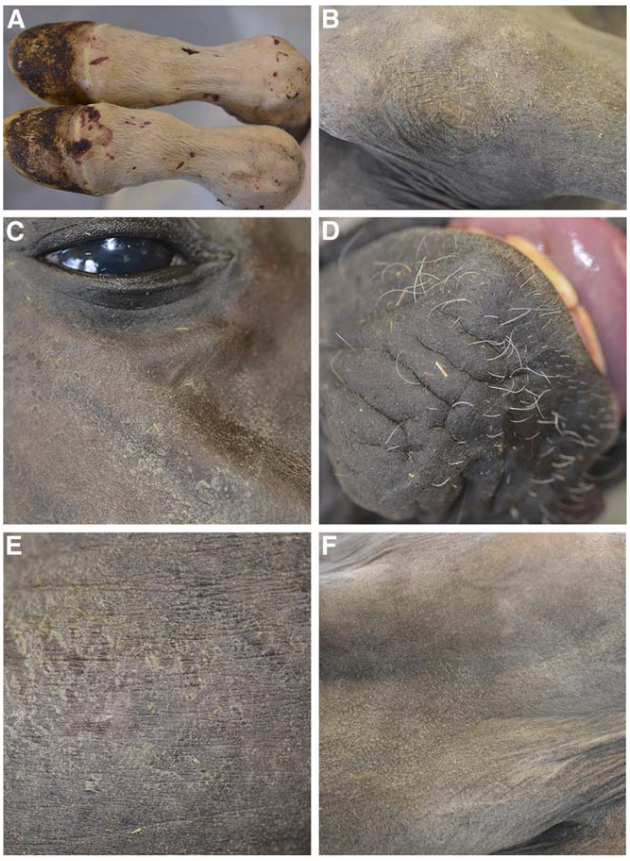


Figure 2 Clinical findings in a 3-wk-old foal with NFS (case 2). (A) Distal fore limbs showed the presence of sparse hairs and multiple abrasive lesions. (B) Dry and scaly skin on the proximal forelimb (xerosis cutis). (C) Scaly alopecic area on the head and missing eyelashes. (D) Sparse, curly, and abnormally short whiskers on the muzzle. (E) Dry and scaly skin on the flank. (F) Hyperkeratosis in the pectoral region.

Histopathologic findings

We performed a histological examination of skin sections from two NFS-affected horses, an obligate carrier, and a control horse. Histologic examination of skin from the carrier, and the unrelated control, were not noticeably different. Both affected horses revealed severely shortened anagen follicles, in which only remnants of the isthmic portion were present, and the hair bulbs of the majority of the follicles were located at the level of, or only slightly below, the sebaceous glands. The infundibula of the hair follicles were often distorted and filled with excessive infundibular keratin, and, especially at the level of the entrance of the sebaceous duct, with abundant sebum. Furthermore, the follicular lumen was often distended at the level of the sebaceous duct entrance. The epithelium of the infundibula and the rudimentary isthmi were mild to moderately hyperplastic and irregular. The sebaceous glands presented with large empty vacuoles. The hair shafts, if present, were very thin and lacking the normal structure. The dysplastic hair bulbs differed in size, and

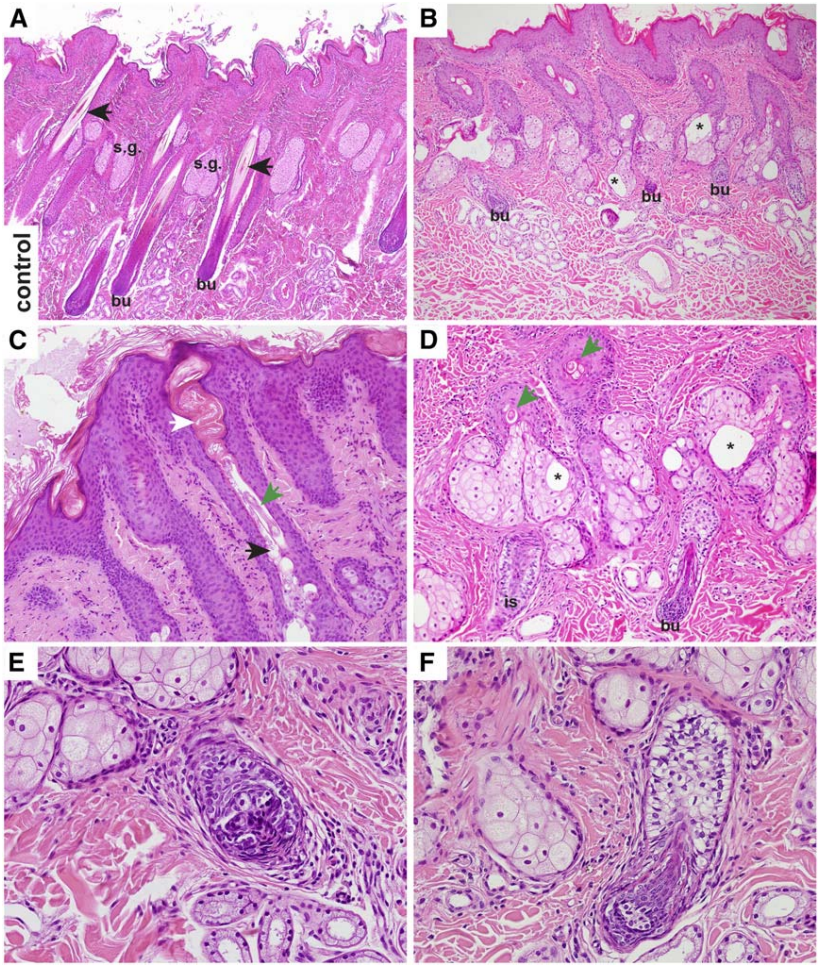


Figure 3 Histopathology of a control and an NFS affected horse (H&E staining). (A) In the control horse, long anagen hair follicles extended into the subcutis, and the bulb (bu) reached far below the level of sebaceous glands (s.g.). The hair follicles contained approximately equally sized, and normally differentiated, hair shafts (arrows). Only little keratin and no sebum were visible in the infundibula. Sebaceous glands had no vacuoles. (B) In the affected horse, the shortened anagen follicles ended at the level of, or just below, the sebaceous glands. There were large empty vacuoles within the sebaceous glands (asterisks). (C) Higher magnification of an infundibulum of an affected horse. The infundibulum contained an increased amount of keratin (white arrow) and sebum (black arrow). The hair shaft present was thin, and had an abnormal structure (green arrow). (D) Higher magnification of the proximal portion of the dysplastic hair follicles. The hair shafts were very thin (green arrows), and the isthmus (is) and bulbar region (bu) appeared only as remnants. Note also the large empty vacuoles within the sebaceous glands. (E) Dysplastic hair bulb in close vicinity to the sebaceous glands of an affected horse. The matrical cells were irregularly arranged, not equal in size, and sometimes vacuolated. (F) Another example of a very small dysplastic bulb in comparison to the adjacent lower isthmus. Note the larger vacuoles within some matrical cells.

the matrical cells of the bulbs were arranged irregularly. Sometimes larger vacuoles were present. The epidermis was moderately hyperplastic, and covered with moderate orthokeratotic laminar to basket weave keratin. In addition, there was a mild perivascular to interstitial lymphocytic infiltrate in the dermis. The histological findings in the hair follicles were compatible with a follicular dysplasia (Figure 3).

Mapping of candidate regions

As the breeders' records indicated a monogenic autosomal recessive mode of inheritance, we used a combined linkage and homozygosity mapping approach to establish the position of the causative variant in the genome (Figure S1). SNP microarray genotypes from three NFS affected half-siblings and their four parents were determined. We performed a parametric linkage analysis assuming a fully penetrant autosomal recessive inheritance model, and obtained 50 linked genomic regions of >1 Mb with a positive LOD score on 26 different chromosomes (Table S3).

We then searched for extended regions of homozygosity in the same three affected horses that had been used for the linkage analysis. Six intervals on chromosomes 7, 18, and 27 fulfilled our homozygosity search criteria (Figure 4 and Table S4). Intersecting the intervals obtained by linkage analysis and homozygosity mapping resulted in four critical intervals located on chromosome 7 and one interval on chromosome 27. Taken together, the four chromosome 7 intervals spanned >15 Mb, and were interrupted only by three small nonhomozygous segments of <1.5 Mb. We therefore conservatively considered the four segments as one whole interval on chromosome 7 for all further analyses. This ~17 Mb interval spanned from position 36,013,146 to position 53,111,271 on chromosome 7. The single 1.2 Mb interval on chromosome 27 spanned from 10,880,351 to 12,070,934.

Whole genome sequencing and identification of the causative variant

To further characterize the identified genomic regions on chromosome 7 and 27, we sequenced the genomes of two affected horses and two obligate carriers at ~20× to 33× coverage. For each sample, we called single nucleotide variants and small indels with respect to the equine reference genome assembly. Since the disorder is only known in the Akhal-Teke breed, we hypothesized that the causative variant should be completely absent in horses of other breeds. We therefore filtered for variants that were present in homozygous state in the affected horses, in heterozygous state in the carriers, and absent or missing in 75 control horses of different breeds.

In the interval on chromosome 27, none of the variants fulfilled our search criteria. In the 17 Mb region on chromosome 7, we found four variants that showed perfect cosegregation with NFS (Table 1). Among these variants, the only nonsynonymous variant was a nonsense variant in the *ST14* gene (c.388G>T) introducing a premature stop codon into

Figure 4 Combined linkage and homozygosity mapping. The 50 linked chromosomal segments >1 Mb are shown in yellow. Homozygous intervals >1 Mb located on chromosomes 7, 18, and 27 are indicated in red. Intersecting the regions obtained by the two methods resulted in critical intervals on chromosomes 7 and 27 (arrows).

exon 4. *ST14* encodes "suppression of tumorigenicity 14," a type II serine protease that is highly expressed in the epidermis, and was reported to be essential for epidermal barrier function, and development of hair follicles (List *et al.* 2002).

We confirmed the presence of the *ST14*:c.388G>T variant by Sanger sequencing on the genomic and transcript level. In addition, we genotyped all available samples from Akhal-Teke horses, as well as 400 horses from various breeds for this variant. As expected, the variant was absent in all tested horses outside of the Akhal-Teke breed. It was homozygous in the five available affected horses, and heterozygous in 10 obligate carriers (Table 2).

The *ST14*:c.388G>T variant introduces a premature stop codon predicted to truncate >80% of the open reading frame (p.Glu130*). We hypothesized that mRNA carrying this variant is likely to be degraded by nonsense-mediated decay. To test this hypothesis, we isolated RNA from peripheral blood mononuclear cells from an obligate carrier. Sanger sequencing of a genomic DNA fragment, and the corresponding cDNA fragment of this heterozygous animal, indicated partial nonsense-mediated decay of the mutant transcript (Figure 5).

DISCUSSION

In this study, we used a combined positional mapping and whole genome sequencing approach to unravel the genetic background of NFS in the Akhal-Teke breed. We identified two critical intervals on chromosomes 7 and 27. The *ST14*:c.388G>T nonsense variant was the only nonsynonymous trait-associated variant in the critical intervals. It was perfectly associated with NFS in all tested Akhal-Teke horses, and completely absent from horses of other breeds. Taken together, the genetic association, the fact that *ST14*:c.388G>T is a nonsense variant, and the knowledge on the *ST14* gene from other species, strongly suggest that this is indeed the causative variant for NFS in Akhal-Teke horses.

Table 1 Variants detected by whole genome resequencing of two cases, two obligate carriers, and 75 controls

Filtering Step	Number of Variants
Variants with genotypes 1/1 in cases and 0/1 in carriers	53,683
NFS-associated variants in whole genome[a]	351
NFS-associated variants in critical interval on chr. 27[a]	0
NFS-associated variants in critical interval on chr. 7[a]	4
NFS-associated nonsynonymous variants in critical intervals[a]	1

[a] "NFS-associated" indicates variants that were homozygous for the alternate allele in the two cases, heterozygous in the two carriers, and homozygous reference (or missing) in the 75 control horses.

The suppressor of tumorigenicity 14 gene (*ST14*) encodes a type II serine protease that was also termed matriptase. It was first described in the context of cancer, but later studied as a protease involved in epithelial, and, in particular, epidermal development (Shi *et al.* 1993; Zhang *et al.* 1998; Miller and List 2013). In *St14* deficient knockout mice, all pups homozygous for the null allele died within 48 hr of birth (List *et al.* 2002). They showed malformed epidermal surfaces with a dry, red, wrinkled, and shiny appearance, and had a lower body weight and smaller size than heterozygous or homozygous wildtype littermates. Abnormalities in stratum corneum development, and loss of both inward and outward barrier functions, were reported. As a consequence, the rate of epithelial fluid loss was significantly accelerated in *St14* deficient mice, resulting in rapid postnatal death. Histological examination showed that vesicular bodies, the major source of intercorneocyte lipids, were not present in transitional cells, which could explain the impaired barrier function. Furthermore, vibrissal hairs were absent in *St14* knockout pups, hair follicles were hypomorphic, and the hair shafts were curved and ingrown, possibly due to the absence of the hair canal (List *et al.* 2002).

In humans, *ST14* variants may cause autosomal recessive congenital ichthyosis 11 (ARCI11, OMIM #602400), which can be further subclassified into autosomal recessive ichthyosis with hypotrichosis (ARIH) and ichthyosis, follicular atrophoderma, hypotrichosis, and hypohidrosis (IFAH). In three members of a consanguineous Israeli-Arab family, ARIH was reported to be caused by the p.Gly827Arg missense variant in the *ST14* gene. The patients showed diffuse and generalized hypotrichosis, diffuse scaling and curly, fragile, sparse slowly growing hair, photophobia, and mild abnormalities of the teeth. Similar to horses with NFS, the hair shafts of the human patients were irregular. Further examination of skin samples of a patient revealed that intact corneodesmosomes were present in the upper layers of the stratum corneum, which indicated an impaired degradation of corneodesmosomes, and an involvement of ST14 in desquamation (Basel-Vanagaite *et al.* 2007). A female patient with ARIH caused by the

Table 2 Association of the ST14:c.388G>T variant with NFS

Genotype	G/G	G/T	T/T
Affected horses (*N* = 5)	0	0	5
Obligate carriers (*N* = 10)	0	10	0
Remaining Akhal-Teke horses (*N* = 191)	165	26	0
Control horses from other breeds (*N* = 400)	400	0	0

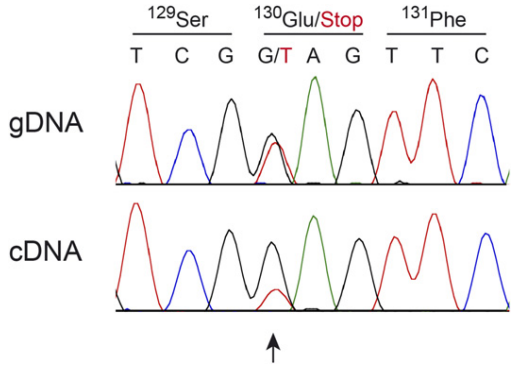


Figure 5 Nonsense-mediated decay of mRNA containing the *ST14*:c.388G>T variant. A fragment containing the variant was amplified from genomic DNA (gDNA), or cDNA of a heterozygous carrier animal and analyzed by Sanger sequencing. The relative peak heights at the heterozygous position differ between sequences amplified from gDNA or cDNA. While the genomic sequence shows the expected 1:1 ratio between the alleles, the mutant allele is underrepresented at the cDNA (transcript) level.

variant c.3G>A, which shifts the translation initiation, did not show any tooth abnormalities, but sparse eyebrows and eyelashes (Avrahami *et al.* 2008). While the ARIH variants were all *ST14* missense variants, one of the IFAH patients had a splice site variant (c.2269+1G>A). In another IFAH patient, a frame shift deletion (c.2034delG) was described. Both variants were predicted to result in premature stop codons. Functional analyses on keratinocytes of the patient with the deletion variant revealed that ST14 was completely missing, profilaggrin processing was altered, and prostasin, a GPI-anchored serine protease with a role in the same proteolytic cascade as ST14, was found only in its inactive form (Alef *et al.* 2009).

Reduced proteolytic activation of prostasin, and reduced processing of profilaggrin were also found in *St14* hypomorphic mice that developed a phenotype comparable to ARIH (List *et al.* 2007). A study on *St14* deficient mice further demonstrated that filaggrin monomers, and the filaggrin S-100 protein that promotes terminal differentiation of keratinocytes were missing, indicating a specific role for ST14 in these processing steps, and in terminal epidermal differentiation (List *et al.* 2003).

Taken together, the *ST14* gene has essential roles in the interfollicular epidermis by contributing to epidermal barrier formation as well as for hair follicle development. The observed phenotype in NFS-affected horses resembles the group of heterogeneous phenotypes caused by *ST14* variants in humans and mice. However, in horses with NFS, the degree of alopecia is more severe than in human patients, whereas the ichthyosis is less pronounced, and suggested mainly by the clinical picture. It is not fully clear why NFS affected foals have such a short life expectancy. While none of the known NFS foals so far became older than 3 yr, there also was huge variability, with some foals dying at a few weeks of age, and others surviving at least for >2 yr. Given the variability in life span, we consider the hydrocephalus, as well as the heart and immune system changes, in our case 2 as coincidental findings unrelated to NFS, but this should be reinvestigated if further NFS-affected foals become available for pathologic examination. Further research will be necessary to characterize the pathologies leading to premature death in these horses.

In conclusion, we provide an initial scientific description of the phenotype of NFS, a trait that has been segregating in the Akhal-Teke breed for >75 yr. Furthermore, we identified the *ST14*:c.388G>T nonsense variant as most likely underlying genetic defect. These findings thus provide a potential large animal model for related human diseases, and allow genetic testing to avoid the nonintentional breeding of NFS-affected foals.

ACKNOWLEDGMENTS

The authors thank all owners who donated samples, photos, and data to the study. Jens Tetens and Georg Thaller are gratefully acknowledged for sharing horse genome sequence data. The authors also thank Nathalie Besuchet, Muriel Fragnière, and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform of the University of Bern is acknowledged for performing the whole genome resequencing experiments, and the Vital-IT high-performance computing center of the Swiss Institute of Bioinformatics for performing computationally-intensive tasks (http://www.vital-it.ch/). This study was financed in part by grants from the World Wide Web WWW Foundation and the Swiss National Science Foundation (CRSII3_160738/1).

LITERATURE CITED

Abecasis, G. R., S. S. Cherny, W. O. Cookson, and L. R. Cardon, 2002 Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat. Genet. 30: 97–101.

Abitbol, M., P. Bossé, A. Thomas, and L. Tiret, 2015 A deletion in *FOXN1* is associated with a syndrome characterized by congenital hypotrichosis and short life expectancy in Birman cats. PLoS One 10: e0120668.

Alef, T., S. Torres, I. Hausser, D. Metze, U. Türsen *et al.*, 2009 Ichthyosis, follicular atrophoderma, and hypotrichosis caused by mutations in *ST14* is associated with impaired profilaggrin processing. J. Invest. Dermatol. 129: 862–869.

Avrahami, L., S. Maas, M. Pasmanik-Chor, L. Rainshtein, N. Magal *et al.*, 2008 Autosomal recessive ichthyosis with hypotrichosis syndrome: further delineation of the phenotype. Clin. Genet. 74: 47–53.

Basel-Vanagaite, L., R. Attia, A. Ishida-Yamamoto, L. Rainshtein, D. Ben Amitai *et al.*, 2007 Autosomal recessive ichthyosis with hypotrichosis caused by a mutation in *ST14*, encoding type II transmembrane serine protease matriptase. Am. J. Hum. Genet. 80: 467–477.

Cadieu, E., M. W. Neff, P. Quignon, K. Walsh, K. Chase *et al.*, 2009 Coat variation in the domestic dog is governed by variants in three genes. Science 326: 150–153.

Drögemüller, C., E. K. Karlsson, M. K. Hytönen, M. Perloski, G. Dolf *et al.*, 2008 A mutation in hairless dogs implicates *FOXI3* in ectodermal development. Science 321: 1462.

Gandolfi, B., C. A. Outerbridge, L. G. Beresford, J. A. Myers, M. Pimentel *et al.*, 2010 The naked truth: Sphynx and Devon Rex cat breed mutations in *KRT71*. Mamm. Genome 21: 509–515.

Gandolfi, B., H. Alhaddad, S. E. Joslin, R. Khan, S. Filler *et al.*, 2013 A splice variant in *KRT71* is associated with curly coat phenotype of Selkirk Rex cats. Sci. Rep. 3: 2000.

Kuznetcova, Y., 2006 Again about sphynxes. Akhal-Teke Inform. 2007: 164–165 (in Russian).

Kuznetcova, Y., M. Kozyreva, and N. Aleksandrova, 2006 The Stavropol Sphynx. Akhal-Teke Inform. 2006: 144–147 (in Russian).

Leisson, K., K. Alev, P. Kaasik, . Jaakma, and T. Seene, 2011 Myosin heavy chain pattern in the Akhal-Teke horses. Animal 5: 658–662.

Leroy, G., and R. Baumung, 2011 Mating practices and the dissemination of genetic disorders in domestic animals, based on the example of dog breeding. Anim. Genet. 42: 66–74.

List, K., C. C. Haudenschild, R. Szabo, W. Chen, S. M. Wahl *et al.*, 2002 Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. Oncogene 21: 3765–3779.

List, K., R. Szabo, P. W. Wertz, J. Segre, C. C. Haudenschild *et al.*, 2003 Loss of proteolytically processed filaggrin caused by epidermal deletion of Matriptase/MT-SP1. J. Cell Biol. 163: 901–910.

List, K., B. Currie, T. C. Scharschmidt, R. Szabo, J. Shireman *et al.*, 2007 Autosomal ichthyosis with hypotrichosis syndrome displays low matriptase proteolytic activity and is phenocopied in ST14 hypomorphic mice. J. Biol. Chem. 282: 36714–36723.

Miller, G. S., and K. List, 2013 The matriptase-prostasin proteolytic cascade in epithelial development and pathology. Cell Tissue Res. 351: 245–253.

Murgiano, L., D. P. Waluk, R. Towers, N. Wiedemar, J. Dietrich *et al.*, 2016 An intronic *MBTPS2* variant results in a splicing defect in horses with brindle coat texture. G3 6: 2963–2970.

Nehls, M., D. Pfeifer, M. Schorpp, H. Hedrich, and T. Boehm, 1994 New member of the winged-helix protein family disrupted in mouse and rat nude mutations. Nature 372: 103–107.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira *et al.*, 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81: 559–575.

Shi, Y. E., J. Torri, L. Yieh, A. Wellstein, M. E. Lippman *et al.*, 1993 Identification and characterization of a novel matrix-degrading protease from hormone-dependent human breast cancer cells. Cancer Res. 53: 1409–1415.

Zhang, Y., X. Cai, B. Schlegelberger, and S. Zheng, 1998 Assignment of human putative tumor suppressor genes ST13 (alias SNC6) and ST14 (alias SNC19) to human chromosome bands 22q13 and 11q24/ q25 by in situ hybridization. Cytogenet. Cell Genet. 83: 56–57.

Communicating editor: D. L. Bannasch

A curated catalog of canine and equine keratin genes

Journal:	PLoS One
Manuscript status:	published
Contributions:	Data curation, formal analysis, investigation, original draft, review and editing of manuscript, Figures 1, 3, 4

PLOS ONE

RESEARCH ARTICLE

A curated catalog of canine and equine keratin genes

Pierre Balmer[1,2☯], Anina Bauer[2,3☯], Shashikant Pujar[4], Kelly M. McGarvey[4], Monika Welle[2,5], Arnaud Galichet[2,6], Eliane J. Müller[2,5,6,7], Kim D. Pruitt[4], Tosso Leeb[2,3], Vidhya Jagannathan[2,3]*

1 Division of Clinical Dermatology, Department of Clinical Veterinary Medicine, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **2** Dermfocus, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **3** Institute of Genetics, Vetsuisse Faculty, University of Bern,Bern, Switzerland, **4** National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, United States of America, **5** Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, Bern, Switzerland, **6** Department of Clinical Research, Molecular Dermatology and Stem Cell Research, University of Bern, Bern, Switzerland, **7** Clinic for Dermatology, Inselspital, Bern University Hospital, Bern, Switzerland

☯ These authors contributed equally to this work.
* vidhya.jagannathan@vetsuisse.unibe.ch

Citation: Balmer P, Bauer A, Pujar S, McGarvey KM, Welle M, Galichet A, et al. (2017) A curated catalog of canine and equine keratin genes. PLoS ONE 12(8): e0180359. https://doi.org/10.1371/journal.pone.0180359

Editor: Claire Wade, University of Sydney Faculty of Veterinary Science, AUSTRALIA

Received: March 2, 2017

Accepted: June 14, 2017

Published: August 28, 2017

Data Availability Statement: The RNA-seq data are available at the European Nucleotide Archive (ENA) under accessions PRJEB14109 (ERS1324262, ERS1165145, ERS1165144), PRJNA78827 and PRJEB3095 (ERS149612). The newly determined genomic sequences are available under accessions LT576418 and LT576419. Complete representations of the equine KRT78 and KRT85 genes, which are partly based on the new primary accessions, are available under accessions LT576420 and LT576421.

Abstract

Keratins represent a large protein family with essential structural and functional roles in epithelial cells of skin, hair follicles, and other organs. During evolution the genes encoding keratins have undergone multiple rounds of duplication and humans have two clusters with a total of 55 functional keratin genes in their genomes. Due to the high similarity between different keratin paralogs and species-specific differences in gene content, the currently available keratin gene annotation in species with draft genome assemblies such as dog and horse is still imperfect. We compared the National Center for Biotechnology Information (NCBI) (dog annotation release 103, horse annotation release 101) and Ensembl (release 87) gene predictions for the canine and equine keratin gene clusters to RNA-seq data that were generated from adult skin of five dogs and two horses and from adult hair follicle tissue of one dog. Taking into consideration the knowledge on the conserved exon/intron structure of keratin genes, we annotated 61 putatively functional keratin genes in both the dog and horse, respectively. Subsequently, curators in the RefSeq group at NCBI reviewed their annotation of keratin genes in the dog and horse genomes (Annotation Release 104 and Annotation Release 102, respectively) and updated annotation and gene nomenclature of several keratin genes. The updates are now available in the NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene).

Introduction

Keratins are intermediate filament proteins of the epithelial cytoskeleton. They are expressed in a cell-, tissue- and differentiation-dependent manner in stratified (e.g. epidermis and cornea) and simple (liver, pancreas and intestine) epithelia as well as in skin appendages such a hairs and nails [1,2]. As structural proteins, keratins provide mechanical stability to maintain

Funding: This study was funded by a grant from the Swiss National Science Foundation (CRSII3_160738 / 1). Work carried out by NCBI authors was funded by the Intramural Research Program of the NIH, National Library of Medicine.

Competing interests: The authors have declared that no competing interests exist.

epithelial integrity and barrier function [3]. This is best exemplified in the epidermis. Keratins represent a major protein fraction of the keratinocytes, the main cell type in the epidermis. In the different layers of the epidermis, keratinocytes differentially express specific keratins, which correlate with their differentiation stage and contribute to the integrity of the epidermis through interaction with both cell-matrix and intercellular adhesion complexes [4,5]. Additionally, keratins are also involved in the regulation of cellular processes such as embryonic development as well as cell motility, proliferation and death by modulating signal molecule activity [6–11]. Furthermore, keratins were recently found to be major regulators of cellular stiffness, a key parameter in cancer development. Thus, the correct expression of specific keratin genes is essential for normal skin homeostasis [1].

Based on their isoelectric point, keratins can be grouped into type I (acidic) and type II (neutral or basic) keratins. Type I and type II keratins form heterodimers, which further assemble into intermediate filaments [12]. Keratins can also be grouped according to their specific expression pattern: 37 human keratins are specifically expressed in epithelia, while 17 are considered hair keratins. However, nine of the 37 epithelial keratins are also restricted to different compartments of the hair follicles [13].

The latest annotation of the human genome contains 55 functional keratin genes and so far, for most of them only a single transcript is known [14–16]. The human keratin genes are clustered at two different chromosomal locations. The first cluster on chromosome 17q21 contains all type I keratin encoding genes except for *KRT18*. The type II keratin encoding genes and *KRT18* are clustered on chromosome 12q13 [17]. Consistent with their important cellular functions, genetic variants in the keratin genes may cause abnormalities in skin, nails, hair and mucosa. Different genetic variants in at least 18 keratin genes have been found to be causative for human genodermatoses, hereditary diseases of the skin [18,19].

There is a high conservation of the keratin genes in mammals with respect to their organization in the genome, but also with respect to their conserved exon/intron structure suggesting multiple duplication events from an ancestral gene during evolution [20]. The mouse has 54 functional keratin genes, organized in two clusters on chromosomes 11 and 15, similar as in humans [17]. Dogs and horses have draft genome assemblies of relatively high quality, but their annotations are almost exclusively based on computational methods [21–24]. The high similarity between the numerous keratin genes as well as sequencing errors and gaps in the reference genome assembly make these predictions error prone. In the current dog and horse annotations there are examples, where exons from different keratin genes have been erroneously merged into computer-predicted keratin transcripts (e.g. Ensembl transcript ENSECAT00000023303 is composed of one exon of *KRT78* and six exons of *KRT8*).

Reliable keratin gene annotations are a requirement for research on heritable skin or hair disorders [25]. Therefore, the aim of this study was to present a curated catalog of canine and equine keratin genes. We based the curated annotation on evolutionary conserved structural features and homology to the corresponding human keratin genes complemented by canine and equine RNA-seq data. We also determined the genomic sequence of some critical gaps in the horse reference genome assembly to facilitate the correct annotation of keratin genes. NCBI RefSeq [26] curators then used our curated annotation and suggested nomenclature to guide updates to their annotation where possible.

Methods

Ethics statement

Animal work consisted of collecting blood samples and skin biopsies of privately owned dogs and horses with owners' consent. All animal experiments were performed in accordance with

the relevant local guidelines and approved by the committee for animal experiments of the canton of Bern (permits no. BE31/13 and BE75/16).

Analyzed dataset, genome assemblies, and existing annotations

The human reference genome GRCh38.p2 assembly and the NCBI annotation release 107 were used for this investigation [14]. For the dog, we used the CanFam 3.1 assembly and the NCBI annotation release 103 (RefSeq release 75). For the horse, we used the EquCab 2.0 assembly and the NCBI annotation release 101 (RefSeq release 75). We also compared our data to the canine and equine annotation from Ensembl (release 87 for both species).

NCBI annotation updates

This manuscript reflects a static comparison to dog annotation release 103 and horse annotation release 101. Since our initial analysis, the NCBI annotation was updated to dog annotation release 104 and horse annotation release 102 by routine employment of NCBI's Eukaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/process/). These pipeline updates resulted in improved agreement with our annotation, as several NCBI gene models that previously differed from our annotation were updated to models that match our annotation. Furthermore, RefSeq curators collaborated to perform a manual review of keratin gene annotation in dog and horse (dog annotation release 104 and horse annotation release 102), which resulted in additional updates. These updates include replacement of model RefSeqs (XM_, XP_, XR_) with known RefSeqs (NM_, NP_, NR_) and updates to gene names. NCBI has adopted most of our suggested nomenclature. For 21 of the dog keratin genes and 28 of the horse keratin genes the symbol was updated in RefSeq. (We have attached NIH publishing agreement as S3 File).

Targeted sequencing of equine keratin gene fragments

We determined the sequence (accessions LT576418 and LT576419) of two genomic fragments from *KRT78* and *KRT85* that were missing from the equine reference genome assembly (chr6:69,933,880–69,934,077 and chr6:69459932–69460612 respectively). DNA from equine EDTA blood (sample FM2644 derived from a Franches-Montagnes horse) was isolated using the Nucleon Bacc2 kit (GE Healthcare Life Sciences) and these regions were PCR amplified using primers `KRT78_F, TAAAGGAAAGGGTCCTGCAA` and `KRT78_R, GAGCGGGTCTCC AGAGATG` or `KRT85_F, TCTTCTTCTTGAAGCTTGACCTG` and `KRT85_R, ACACCCAGC ACAGGCAGAC`. PCR products were treated with shrimp alkaline phosphatase and exonuclease I and then sequenced on both strands using ABI v3.1 BigDye chemistry and an ABI3730 capillary sequencer.

RNA-seq

We isolated total RNA from skin biopsies of 3 dogs (DS032, DS042, LA1666 (PRJEB14110 and PRJEB14109)), hair follicle tissue from 1 dog (PRJEB14110), and a skin biopsy from 1 horse (UKH004) (PRJEB12979) using the RNeasy Fibrous Tissue Mini kit (Qiagen). Prior to RNA extraction the tissue was mechanically disrupted using the TissueRuptor device (Qiagen). The RNA samples were transformed into illumina TruSeq libraries and 2 x 150 bp sequencing reads were obtained on a HiSeq3000 instrument (illumina) at the Next Generation Sequencing Platform, University of Bern. The reads were filtered for low quality bases using a Phred quality score threshold of 15 for each base and reads longer than 50 bases were retained. The star aligner [27] was used to map the quality filtered reads to Ensembl CanFam3.1 reference using

recommended parameters '—outFilterType BySJout' and '—outFilterMultimapNmax 20'. We additionally used publicly available illumine skin RNA-seq datasets from the skin of two dogs, a Beagle and a dog of unspecified breed [22] as well as from the skin biopsy of one domestic horse showing Leopard complex spotting (EBI accessions PRJNA78827 and PRJEB3095). The reads of the downloaded RNA-seq data sets were also filtered and mapped to the reference genome in the same manner as described above. Exon level coverage was calculated using BedTools [28].

Whole genome sequence data

83 paired end whole genome sequences from 28 diverse breeds of horses were mapped to the reference genome (NCBI) EquCab2.0 using BWA-MEM (version 0.7.10). The aligned reads were further coordinate sorted with samtools 1.2 [29] and duplicate marked with picard tools (http://broadinstitute.github.io/picard/). The whole genome datasets are available publicly at the European Nucleotide Archive (ENA; study accessions PRJEB14779, PRJEB5942, PRJEB10098, PRJEB12984, PRJEB12979, PRJEB9269, PRJEB9267, PRJEB9139, PRJEB14651).

Comparison of RefSeq transcripts to RNA-seq data

Keratin gene clusters in the canine and equine genomes were identified by searching the NCBI Gene database for keratin genes corresponding to the first and last keratin gene of the orthologous human clusters. The identified chromosomal regions were inspected using Genome Data Viewer [30]. All the NCBI RefSeq transcripts in the cluster regions of the human, dog and horse genomes were downloaded from NCBI (S1 Table). If a keyword search with the human gene symbol did not lead to the corresponding horse or dog keratin gene, the genomic sequence of the human keratin gene of interest was used as query in a BLASTN search against the genome of the respective species. BLASTN hits and the relative position of the potential dog or horse keratin gene within the cluster compared to the relative position within the human keratin gene cluster were then used to identify putatively orthologous keratin genes.

Untranslated regions (UTRs), start and stop codons of the coding sequence as well as exon-intron boundaries of the RefSeq transcripts were visually compared to RNA-seq data from dog skin and dog hair follicles as well as from horse skin using the Integrative Genome Viewer (IGV) [31]. In case the RNA-Seq data did not support a coordinate of the RefSeq transcript annotation and if there was clear evidence from the RNA-Seq data for different exon coordinates, the RNA-seq supported exon coordinates were adopted for the curated annotation, which in case of equine *KRT9* and *KRT41* identified frameshifts of the conserved open reading frames in the genomic reference sequence.

Results

Genomic organization of keratin gene clusters

Type I keratin genes except for *KRT18* are clustered on human chromosome 17 (HSA 17), the corresponding gene clusters in dogs and horses are located on chromosomes 9 and 11 (CFA 9 and ECA 11). Compared to the human cluster, the keratin type I gene cluster is in reverse orientation in both dog and horse (Fig 1; S1 Table; S2 Table).

There are 28 functional keratin genes and four pseudogenes (gene names with suffix P) in the human keratin type I gene cluster on HSA 17. Many 1:1 orthology relationships between the human, canine and equine genes exist and the synteny is largely conserved. However, in dogs, the homologs of two human pseudogenes appear to be still active and functional (*KRT41P* and *KRT42P*). Transcription of both genes is supported by the available RNA-seq data (average

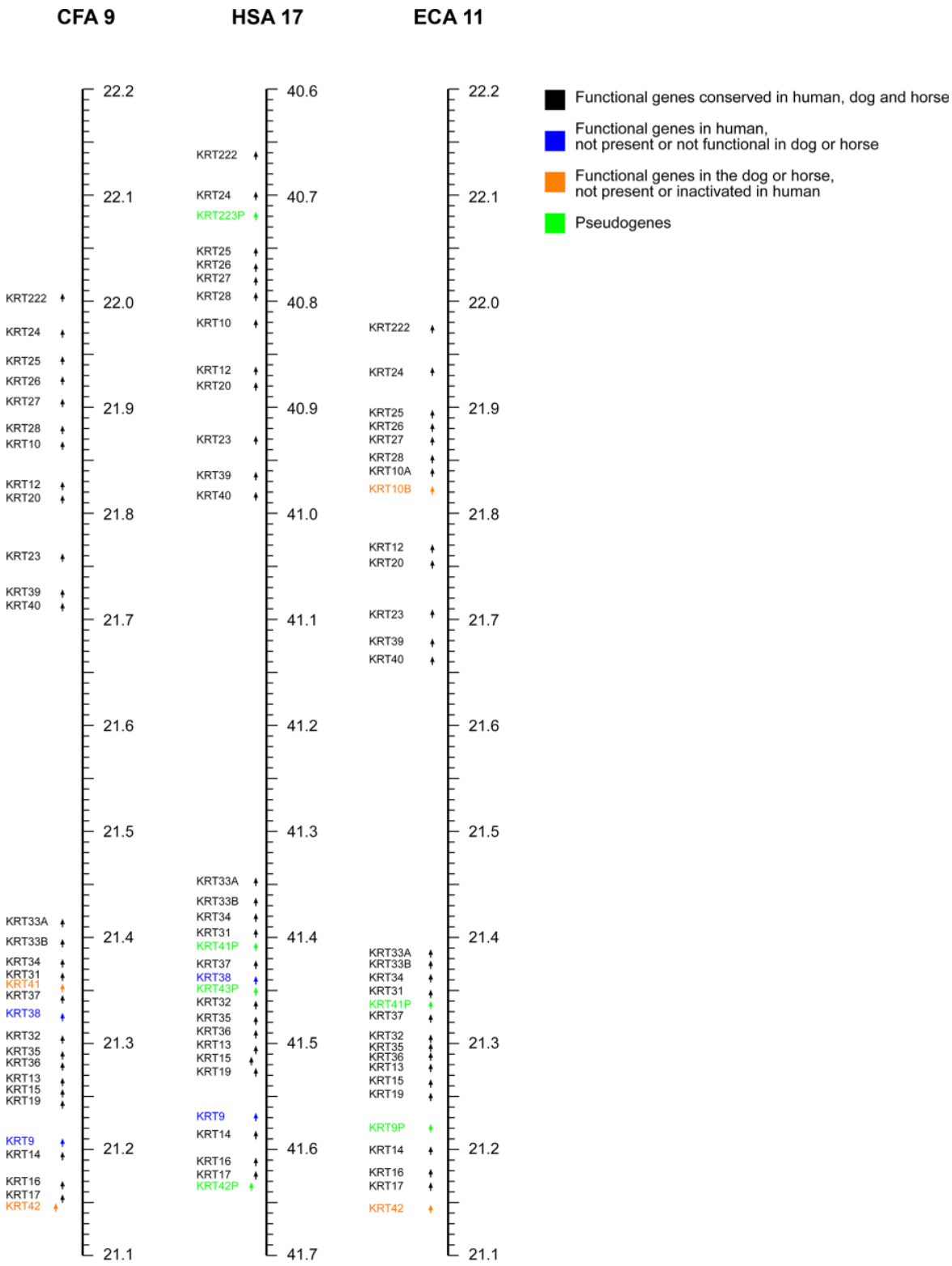


Fig 1. Comparative map of the keratin type I gene cluster. Type I keratin genes except KRT18 in the dog, human and horse genomes. Arrows indicate the orientation of the genes. Note that CFA 9 and ECA 11 are represented in reverse orientation with decreasing coordinates from top to bottom. The genes' nomenclature in the figure corresponds to our updated nomenclature for keratin genes (see Methods). S3 Table details the correspondence of gene symbols to NCBI Gene IDs and RefSeq IDs.

https://doi.org/10.1371/journal.pone.0180359.g001

exon coverage of 21x and 161x respectively). We suggest the gene symbols *KRT41* and *KRT42* for the corresponding canine genes. Homologs for the other two human pseudogenes *KRT43P* and *KRT223P* are missing in the dog reference genome assembly. Thus the canine keratin type I gene cluster most likely contains 30 functional genes.

In the horse there has been a duplication event of the *KRT10* gene with respect to humans and dogs (S1 Fig). We suggest the gene symbols *KRT10A* and *KRT10B* for the corresponding equine paralogs. On the other hand, there is no equine ortholog of the human and canine *KRT38* gene in the current horse genome assembly. *KRT42* is expressed, contains an intact open reading frame, and thus appears to be functional in horse, similar to dogs and differing from humans. The equine homolog of the human *KRT9* genes contains a 1 bp deletion in exon 6 compared to humans and most other mammals. The deletion leads to a frameshift and an early premature stop codon only two codons after the deletion. The genome reference and whole genome sequences from 83 genetically diverse horses are all in agreement without any evidence for intraspecies variation at the position of this deletion. RNA-seq also confirmed the presence of the 1 bp deletion in the equine homolog of the *KRT9* gene (Fig 2, S2 Fig). Thus the equine homolog of *KRT9* is most likely no longer functional. We therefore suggest the gene symbol *KRT9P* for the equine pseudogene. The equine homolog of human *KRT41P* contains a frameshift in the first exon and is thus also likely a pseudogene, for which we suggest the gene symbol *KRT41P*. There were no equine homologues of the remaining two human pseudogenes *KRT43P* and *KRT223P* and we did not find any evidence for transcribed pseudogenes or functional genes in these regions in our RNA-seq data. The orthologs of these pseudogenes might be missing from the horse reference genome assembly, similar to the situation in dogs. Thus, the equine keratin type I gene cluster most likely contains 28 functional genes and two pseudogenes (Table 1).

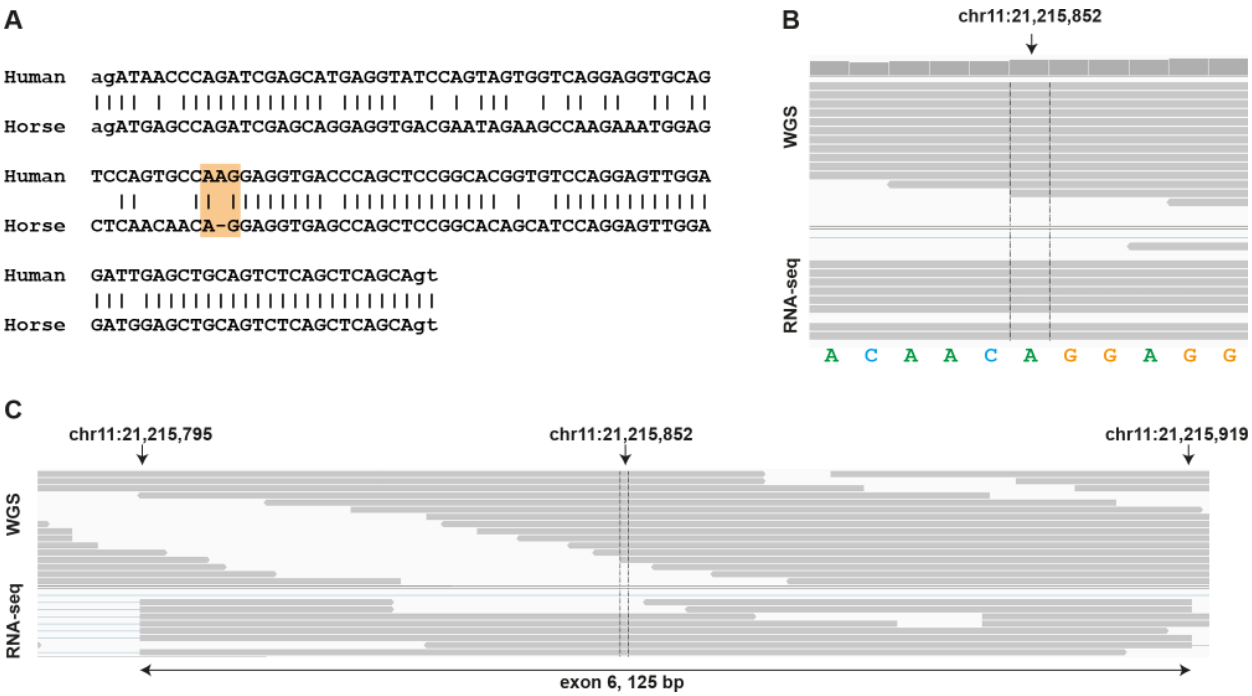


Fig 2. Frameshift deletion in exon 6 of the equine KRT9 **gene.** A) Exon 6 alignment of equine KRT9P against the human ortholog KRT9. The yellow block shows the deletion of a single base in the horse exon leading to a frameshift. B) Illumina whole genome sequence and RNA-seq alignment confirms that the equine reference sequence does not contain a sequencing error which supports the claim of deletion. C. Expanded alignment region containing the frameshift deletion.

https://doi.org/10.1371/journal.pone.0180359.g002

Table 1. Details of the genes in the keratin type I gene cluster on HSA 17, CFA 9, and ECA 11.

HSA 17 (GRCh38.p2)			CFA 9 (CanFam3.1)			ECA 11 (EquCab 2.0)		
Gene symbol	Exons	Encoded amino acids	Gene symbol[a]	Exons	Encoded amino acids	Gene symbol[a]	Exons	Encoded amino acids
KRT9	8	983	KRT9	8	674	KRT9P	(9)[b]	(P)[b,c]
KRT10	8	584	KRT10	8	568	KRT10A	8	559
-	-	-	-	-	-	KRT10B	8	596
KRT12	8	494	KRT12	8	507	KRT12	8	488
KRT13	8	458	KRT13	8	452	KRT13	8	451
KRT14	8	472	KRT14	8	482	KRT14	8	478
KRT15	8	456	KRT15	8	472	KRT15	8	469
KRT16	8	473	KRT16	8	477	KRT16	8	(472)[b]
KRT17	8	432	KRT17	8	433	KRT17	8	433
KRT19	6	400	KRT19	6	399	KRT19	6	411
KRT20	8	424	KRT20	8	434	KRT20	8	424
KRT23	8	422	KRT23	8	419	KRT23	8	428
KRT24	8	525	KRT24	8	541	KRT24	8	506
KRT25	8	450	KRT25	8	450	KRT25	8	450
KRT26	8	468	KRT26	8	464	KRT26	8	469
KRT27	8	459	KRT27	8	446	KRT27	8	460
KRT28	8	464	KRT28	8	464	KRT28	8	464
KRT31	7	416	KRT31	7	409	KRT31	7	416
KRT32	7	448	KRT32	7	448	KRT32	7	448
KRT33A	7	404	KRT33A	7	404	KRT33A	7	404
KRT33B	7	404	KRT33B	7	407	KRT33B	7	404
KRT34	7	436	KRT34	7	393	KRT34	7	393
KRT35	7	455	KRT35	7	455	KRT35	7	455
KRT36	7	467	KRT36	7	467	KRT36	7	461
KRT37	7	449	KRT37	7	440	KRT37	7	444
KRT38	7	456	KRT38	7	457	-	-	-
KRT39	7	491	KRT39	7	483	KRT39	7	492
KRT40	7	431	KRT40	7	431	KRT40	7	431
KRT41P	P[c]	P[c]	KRT41	7	437	KRT41P	7	(P)[b,c]
KRT42P	P[c]	P[c]	KRT42	10	(492)[b]	KRT42	8	453
KRT43P	P[c]	P[c]	-	-	-	-	-	-
KRT222	6	295	KRT222	6	295	KRT222	6	295
KRT223P	P[c]	P[c]	-	-	-	-	-	-

[a] The gene symbols have been updated by NCBI Refseq curators. A comprehensive listing of gene symbols and NCBI Gene IDs is given in S3 Table.
[b]Data in brackets are of low reliability due to gaps in the genome reference assemblies and/or insufficient coverage in the canine and equine RNA-seq data.
[c]P indicates pseudogenes.

https://doi.org/10.1371/journal.pone.0180359.t001

The keratin type II gene cluster located on HSA 12 in humans can be found on chromosomes 27 (CFA 27) and 6 (ECA 6) in dogs and horses. The canine gene cluster is in reverse orientation compared to human and horse (Fig 3; S1 Table).

There are 27 functional keratin genes and eight pseudogenes in the human keratin type II gene cluster. Similar to the keratin type I gene cluster, there are again many 1:1 orthology relationships between the human, canine and equine genes and the synteny is also largely conserved. In the canine cluster on CFA 27, the homologs of five human pseudogenes (*KRT88P*, *KRT89P*, *KRT90P*, *KRT126P*, *and KRT127P*) contain intact open reading frames and appear to

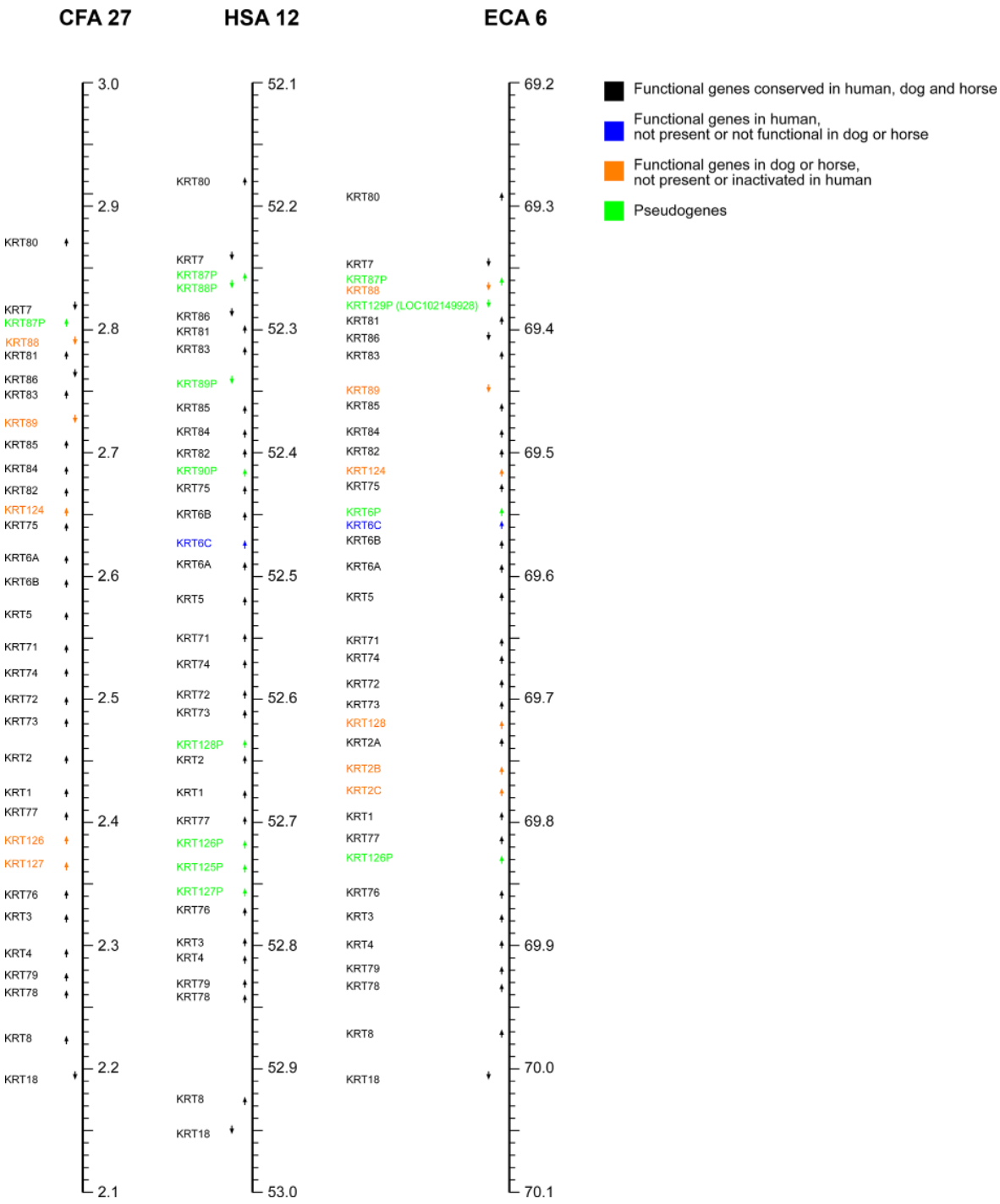


Fig 3. Comparative map of the keratin type II gene cluster. Type II keratin genes and KRT18 in the dog (CanFam3.1), human (GRCh38.p2) and horse genomes (EquCab 2.0). Arrows indicate the orientation of the genes. Note that CFA 27 is represented in reverse orientation with decreasing coordinates from top to bottom. The genes' nomenclature in the figure corresponds to our updated nomenclature for keratin genes (see Methods). S3 Table details the correspondence of gene symbols to NCBI Gene IDs and RefSeq IDs.

https://doi.org/10.1371/journal.pone.0180359.g003

be still functional (average exon level expression of 3.1x, 76.5x, 2.8x, 1024.3x and 71.1x, respectively). S4 Table lists descriptive summary statistics for the predicted canine proteins. Orthologs of the human *KRT6C* gene and the pseudogene *KRT128P* are missing from the dog reference genome assembly. The dog thus most likely has 31 functional genes and one pseudogene in the keratin type II gene cluster.

In the horse there have been duplication events involving the *KRT2* gene with respect to humans and dogs (S3 Fig). We suggest *KRT2A*, *KRT2B* and *KRT2C* as gene symbols for the corresponding equine paralogs. Five of the human pseudogenes appear to be functional in the horse based on intact open reading frames. *KRT6P* is an additional equine pseudogene that is missing from the dog and human reference genome assembly. On the other hand, there is no equine ortholog for the human *KRT125P* and *KRT127P*. The equine keratin gene with Gene ID: 102149928 (*LOC102149928)* is annotated as functional equine keratin gene with five exons. However, there is neither supporting evidence from our adult skin RNA-seq data for the expression of this gene, nor is there an ortholog in the human or dog reference genome assembly. Our suggested nomenclature is *KRT129P*, but this has not been accepted at NCBI due to an underlying gap (chr6:69,382,962–69,383,217) in the horse assembly. As the gene symbol *KRT124* has been proposed in the literature for the equine functional ortholog of human *KRT90P* [32], we suggest the gene symbol *KRT124* for the canine and equine genes (canine Gene ID: 486524, equine Gene ID: 100061458). Thus the equine keratin type II gene cluster most likely contains 33 functional genes and four pseudogenes (Table 2).

We compiled a detailed annotation of canine and equine keratin genes in S2 Table and we provide 61 manually curated canine and equine cDNA sequences in the S1 and S2 Files.

Keratin gene structure

The exon number of functional type I keratin genes varies between 6 and 8 in humans, dogs and horses. The length of exons 2–5 is conserved except for *KRT18*, *KRT23* and *KRT222* in all three species. The same conserved exon structure was found in *KRT42* in dog and horse but not in the human orthologous pseudogene. The length of exon 6 in keratin type I genes is 221 bp ± 3 bp except for *KRT19* and *KRT222*, which only consist of 6 exons in total and *KRT42P* in human.

Type II keratin genes in humans, dogs and horses generally have nine translated exons, except *KRT18*, *KRT74*, *KRT128*, and the canine and equine *KRT88*. The length of exons 3–7 in the type II keratin genes is conserved in all three species with the exception of *KRT8*, *KRT76*, and *KRT74*. The lengths of exons 2 and 8 varied slightly between type II keratin genes (± 3 or 6 bp). Fig 4 illustrates the typical exon/intron organization of keratin genes and S2 Table summarizes the details of all individual exons.

Selected individual keratin genes

During our analyses, we recognized that coding exons of the equine *KRT78* and *KRT85* genes were located in gaps of the reference genome assembly (chr6:69,933,880–69,934,077 and chr6:69,459,932–69,460,612). We performed targeted sequencing of these gaps to provide the correct sequences for exon 4 of *KRT78* (96 bp, accession LT576419) and exon 7 of *KRT85* (221 bp, accession LT576418).

The annotation of canine *KRT77* was not curated due to gaps (chr27:2,412,265–2,412,884) in the reference sequence. The equine *KRT2C* gene also contains a gap probably harboring parts of coding exons (chr6:69,774,035–69,776,390). We provide a partial annotation with corrected exon numbers and inserted the presumed number of missing nucleotides to the compilation of FASTA sequences.

Table 2. Details of the genes in the keratin type II gene cluster on HSA 12, CFA 27, and ECA 6.

HSA 12 (GRCh38.p2)			CFA 27 (CanFam3.1)			ECA 6 (EquCab 2.0)		
Gene symbol	Exons	Encoded amino acids	Gene symbol[a]	Exons	Amino acids	Gene symbol[a]	Exons	Encoded amino acids
KRT1	9	644	KRT1	9	619	KRT1	9	626
KRT2	9	639	KRT2	9	635	KRT2A	9	629
-	-	-	-	-	-	KRT2B	9	618
-	-	-	-	-	-	KRT2C	(9)[b]	(618)[b]
KRT3	9	628	KRT3 (LOC100683401)	9	610	KRT3	9	630
KRT4	9	520	KRT4	9	530	KRT4	9	607
KRT5	9	590	KRT5	9	597	KRT5	9	593
KRT6A	9	564	KRT6A (LOC486523)	9	570	KRT6A	9	574
KRT6B	9	564	KRT6B (LOC100687987)	9	570	KRT6B	9	562
KRT6C	9	564	-	-	-	KRT6C	9	562
-	-	-	-	-	-	KRT6P	P[c]	P[c]
KRT7	9	469	KRT7	9	468	KRT7	9	465
KRT8	9	511	KRT8	9	491	KRT8	9	435
KRT18	7	430	KRT18	7	431	KRT18	6	430
KRT71	9	523	KRT71	9	525	KRT71	9	525
KRT72	9	511	KRT72	9	523	KRT72	9	523
KRT73	9	540	KRT73	9	589	KRT73	9	540
KRT74	9	529	KRT74	9	533	KRT74	(6)[b]	(392)[b]
KRT75	9	551	KRT75	9	551	KRT75	9	549
KRT76	9	638	KRT76	9	648	KRT76	9	640
KRT77	9	578	KRT77	(9)[b]	(561)[b]	KRT77	9	593
KRT78	9	520	KRT78	9	517	KRT78	9	499
KRT79	9	535	KRT79	9	535	KRT79	9	535
KRT80	9	452	KRT80	9	453	KRT80	9	453
KRT81	9	505	KRT81	9	513	KRT81	9	507
KRT82	9	513	KRT82	9	518	KRT82	9	518
KRT83	9	493	KRT83	9	487	KRT83	9	491
KRT84	9	600	KRT84	9	586	KRT84	9	575
KRT85	9	507	KRT85	9	507	KRT85	9	507
KRT86	9	486	KRT86	9	485	KRT86	9	486
KRT87P	P[c]	P[c]	KRT87P	P[c]	P[c]	KRT87P	(8)[b]	(P)[b,c]
KRT88P	P[c]	P[c]	KRT88	(5)[b]	(153)[b]	KRT88	(9)[b]	(483)[b]
KRT89P	P[c]	P[c]	KRT89	9	495	KRT89	9	494
KRT90P	P[c]	P[c]	KRT124	(9)[b]	(531)[b]	KRT124	9	521
KRT125P	P[c]	P[c]	-	-	-	-	-	-
KRT126P	P[c]	P[c]	KRT126	9	630	KRT126P	(n.d.)[b]	(P)[b,c]
KRT127P	P[c]	P[c]	KRT127	9	581	-	-	-
KRT128P	P[c]	P[c]	-	-	-	KRT128	(10)[b]	(514)[b]
-	-	-	-	-	-	KRT129P (LOC102149928)	(P)[b,c]	(P)[b,c]

[a]The gene symbols have been updated by NCBI Refseq curators. A comprehensive listing of gene symbols and NCBI Gene IDs is given in S3 Table.

[b]Data in brackets are of low reliability due to gaps in the genome reference assemblies and/or insufficient coverage in the canine and equine RNA-seq data.

[c]P indicates pseudogenes.

https://doi.org/10.1371/journal.pone.0180359.t002

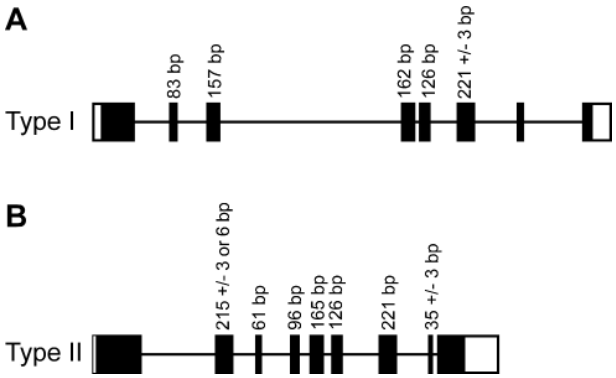


Fig 4. Keratin gene structure. (A) Typical type I keratin gene with eight exons. (B) Typical type II keratin gene with nine exons. The length of conserved exons is given in bp. Untranslated regions are shown as open rectangles.

https://doi.org/10.1371/journal.pone.0180359.g004

Curated keratin annotations

Visual comparison of RNA-seq reads generated from skin of adult dogs and horses with our tentative annotation provided experimental support for 56 out of 62 investigated canine and 47 out of 67 investigated equine keratin genes (S2 Table). For six canine and 20 equine keratin genes, the coverage of the RNA-Seq data was insufficient for a reliable confirmation or curation/revision of the existing annotations.

Comparison of curated keratin annotation with NCBI and Ensembl annotation

We compared our annotation of keratin genes with experimental support to the annotations from NCBI (dog annotation release 103 and horse annotation release 101) and Ensembl (release 87). Roughly half (24 of 51) canine keratin gene annotations were identical between NCBI, Ensembl and our catalog. However, 21 annotations of our catalog were only identical to the NCBI predictions but differed from Ensembl predictions. For the equine keratin genes 21 of our annotated genes were identical with Ensembl and NCBI. Of the remaining equine gene models, 18 were identical with only NCBI annotations and one was only identical to the Ensembl annotation. Five of our canine and eight of our equine curated keratin gene annotations were different from both the NCBI and Ensembl annotation.

Updated NCBI annotation

As a result of updated NCBI annotation via NCBI's Eukaryotic Genome Annotation Pipeline, three dog keratin genes (*KRT4*, *KRT75* and *KRT76*), which were different from our curated annotation in the previous NCBI annotation, are now identical to our annotation. Known (curated) RefSeqs replace model RefSeqs for two of these genes (*KRT4* and *KRT76*). Updates were also made following manual review of the current NCBI annotation of dog and horse keratin genes. These include gene name updates for 21 dog keratin genes and 34 horse keratin genes. Additionally, known RefSeqs replaced model RefSeqs for 19 dog keratin genes and 33 horse keratin genes (S1 Table). The gene type was updated to 'pseudogene' for three horse keratin genes (*KRT87P*, *KRT126P* and *KRT9P*). In the current NCBI annotation, representation of five dog keratin genes and nine horse keratin genes remain different from our curated annotation (S5 Table). However, when excluding curated annotations of low reliability due to insufficient read coverage in the RNA-seq data and keratin genes where no Ensembl annotation is

given, only two canine (*KRT126* and *KRT127*) and six equine (*KRT8*, *KRT18*, *KRT16*, *KRT29*, *KRT78*, *KRT85*) keratin gene annotations differ between the updated NCBI annotation and our curated annotation.

Discussion

We present a curated catalog of both canine and equine keratin genes, which is based on evolutionary conserved features of keratin genes and experimental support from RNA-seq data derived from adult skin. While initial comparison of our curated annotation of dog and horse keratin genes to NCBI annotation revealed several differences, the updated NCBI annotation led to prediction of several new models that were identical to our annotation (S1 Table). Additionally, a manual review of NCBI annotation by RefSeq curators resulted in several updates to gene annotation and gene nomenclature. Differences persist in the annotation of five dog and nine horse keratin genes. Many of the errors in the currently existing keratin gene predictions were due to gaps and errors in the genome reference assemblies. Although automated gene predictions in general reach a very high quality and are essential for scientists to make use of the wealth of publicly available genomic sequence information, this study also shows some of the limitations of automated gene predictions on imperfect genome assemblies.

The most striking example in our study was the *KRT9* gene. Human keratin 9 is mainly expressed in the skin at the soles of feet and the palms of hands, the so-called palmoplantar epidermis. Palmoplantar epidermis needs to withstand the highest mechanical stress on the body and thus has a specialized keratin composition. Keratin 9 is one of the key proteins that provide this extraordinary mechanical strength. Consequently, genetic variants in the *KRT9* gene can lead to palmoplantar keratoderma (OMIM ID:#144200) resulting in the localized fissuring of the palmoplantar epidermis [19,33]. The *KRT9* gene is well conserved between humans and dogs (76.1% identity within the coding sequence of the mRNA). Dogs also have a strong palmoplantar epidermis at their footpads. However, the equine ortholog of the *KRT9* gene contains a 1-nucleotide deletion in exon 6, which we assume causes the functional inactivation of this gene. Automated gene prediction tools infer information from other species and must tolerate sequencing errors, which are inherent to draft genome assemblies. In order to maintain the conserved reading frame this resulted in an NCBI transcript prediction containing an extra nucleotide, which is not present in the genome reference (XM_005614759.1). We now have experimental evidence in the form of Illumina whole genome sequence and RNA-seq data, which both confirm that in this case the equine reference genome assembly was correct and that there is indeed the 1-nucleotide deletion in the equine *KRT9P* pseudogene. In the most recent NCBI annotation release 102, XM_005614759.1 has been since replaced by XM_014736779.1 and the locus type has been updated to pseudogene as a result.

From a biological point of view, it is tempting to speculate that the extreme specialization of the equine epidermis resulting in hoof wall formation and the lack of a palmoplantar epidermis in horses might have led to decreased selection pressure on *KRT9* and eventually the fixation of the frameshift variant. This frameshift deletion in the *KRT9P* pseudogene is present in several equid species (*E. caballus*, *E. przewalski*, *E. asinus*), while the conserved reading frame is maintained in other ungulates such as cattle, goats, and even the perissodactyl southern rhinoceros (S2 Fig).

We think that our curated annotation of canine and equine keratin genes provides a considerable improvement with respect to the currently existing gene predictions. However, we have to caution that we also had to rely on imperfect genome reference assemblies and limited RNA-seq data, mostly restricted to adult skin. Future annotation efforts might benefit from analyzing RNA-seq data of a more comprehensive collection of tissue types and/or developmental stages.

Therefore, we certainly do not claim that our version is perfect and definitive. Further research efforts will be necessary to eventually produce the complete annotation of the vastly complex keratin gene family in dogs and horses.

Conclusions

The study of keratin genes in dog and horse is difficult because of the high duplication rate among the genes which can lead to misassemblies in the reference genomes. The manual annotation of keratin gene clusters enabled not only identification of new pseudogenes in the keratin clusters of both dog and horse, but also the identification of missing or functional canine and equine homologs of human pseudogenes. We provide a curated annotation and representative transcript sequences for canine and equine keratin genes. Correct transcript sequences are essential for many modern high-throughput technologies that rely on the bioinformatic processing of large datasets. This analysis also shows that manual intervention is still critical for the annotation of complex gene families and the production of a good reference gene set for these gene families. This work represents collaboration between a research group and a public annotation database, which resulted in improvement of annotation of a gene family in the dog and horse genomes. Such collaborations provide an opportunity for researchers and annotation groups to work together to represent specific genes or gene families that may not be accurately annotated solely by means of an automated pipeline. The updated NCBI Refseq entries will be incorporated as supporting evidence during future annotation releases by Ensembl. Many instances of low confidence annotation are in regions where the reference assembly is of poor quality, underlining the importance of a high-quality genome assembly as a prerequisite for accurate annotation of genes.

Supporting information

S1 Fig. Comparative analysis of the KRT10 genome region. (A) Dot plot of the human region containing the *KRT28*, *KRT10*, and *KRT12* genes (chr17:40,794,137–40,872,876) against the corresponding dog region (chr9:21,823,976–21,910,085). Human and dog showed a well conserved synteny in this region. (B) Dot plot of the human region against the horse region (chr11:21,764,167–21,862,851). In the horse, a duplication event gave rise to *KRT10A* and *KRT10B* paralogs. (C) The horse-specific duplication also became apparent in the horse vs dog dot plot. Dot plots were generated with a word size of 10 and the software GEPARD. (PDF)
(PDF)

S2 Fig. Multispecies alignment of exon 5 of the human KRT9 gene. (A) This exon has a conserved length of 126 nucleotides (42 codons) in many mammalian species with high quality genome reference assemblies. It corresponds to exon 6 of the equine *KRT9P* pseudogene. In the horse, the Przewalski horse, and the donkey, the gene contains a 1 nt deletion (highlighted in grey) that leads to a frameshift and an early premature stop codon (highlighted in yellow), which truncates ~40% of the conserved open reading frame. As this frameshift deletion occurs in several equid species, it most likely arose during the early evolution of the equid family. The accessions and coordinates of the genomic sequences are given beneath the alignment. (B) Multispecies alignment of the translated amino acid sequences in one letter abbreviations.
(PDF)
(PDF)

S3 Fig. Comparative analysis of the KRT2 genome region. (A) Dot plot of the human region containing the *KRT73*, *KRT128P*, *KRT2*, and *KRT1* genes (chr12:52,607,570–52,680,407) against the corresponding dog region (chr27:2,422,150–2,488,436). Human and dog showed a

well conserved synteny in this region. (B) Dot plot of the human region against the horse region (chr6:69,698,571–69,796,491). In the horse, several duplication events gave rise to the *KRT2A*, *KRT2B*, and *KRT2C* paralogs. The support for the functional status of the equine *KRT2C* and *KRT128* genes was weak and their annotations should be considered of low confidence. (C) The horse-specific amplification also became apparent in the horse vs dog dot plot. Dot plots were generated with a word size of 10 and the software GEPARD. (PDF)
(PDF)

S1 File. FASTA-file containing 60 curated canine keratin transcript sequences. The file is lacking a sequence for canine *KRT77* due to the low reliability of the current annotation.
(TXT)

S2 File. FASTA-file containing 61 curated equine keratin transcript sequences. The file is lacking a sequence for equine *KRT74* due to the low reliability of the current annotation.
(TXT)

S3 File. NIH publishing agreement & manuscript cover sheet.
(PDF)

S1 Table. Compilation of RefSeqs for canine and equine keratin genes from annotation release 104 for dogs and 102 for horses.
(XLSX)

S2 Table. Comparative curated annotation of human, canine and equine keratin genes.
(XLSX)

S3 Table. Nomenclature (gene symbols) of human, canine and equine keratin genes.
(XLSX)

S4 Table. Summary of keratins encoded by genes that appear to be functional in dogs or horses, but inactivated or absent in the human genome.
(XLSX)

S5 Table. Genes with remaining differences between NCBI annotation and our manual curation.
(XLSX)

Acknowledgments

The authors are grateful to Nathalie Besuchet Schmutz, Muriel Fragnière, and Sabrina Schenk for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing whole genome sequencing experiments, and the Interfaculty Bioinformatics Unit of the University of Bern for providing computational infrastructure.

Author Contributions

Conceptualization: Eliane J. Müller, Tosso Leeb.

Data curation: Pierre Balmer, Anina Bauer, Shashikant Pujar, Kelly M. McGarvey, Arnaud Galichet, Kim D. Pruitt.

Formal analysis: Pierre Balmer, Anina Bauer, Shashikant Pujar, Kelly M. McGarvey, Kim D. Pruitt, Tosso Leeb, Vidhya Jagannathan.

Investigation: Pierre Balmer, Anina Bauer.

Resources: Monika Welle, Arnaud Galichet, Eliane J. Müller.

Software: Vidhya Jagannathan.

Visualization: Tosso Leeb.

Writing – original draft: Pierre Balmer, Anina Bauer.

Writing – review & editing: Pierre Balmer, Anina Bauer, Shashikant Pujar, Kelly M. McGarvey, Arnaud Galichet, Eliane J. Müller, Kim D. Pruitt, Tosso Leeb, Vidhya Jagannathan.

References

1. Loschke F, Seltmann K, Bouameur JE, Magin TM. Regulation of keratin network organization. Curr Opin Cell Biol. 2015; 32: 56–64. https://doi.org/10.1016/j.ceb.2014.12.006 PMID: 25594948
2. Hutton E, Paladini RD, Yu QC, Yen M, Coulombe PA, Fuchs E. Functional differences between keratins of stratified and simple epithelia. J Cell Biol. 1998; 143: 487–499. PMID: 9786957
3. Coulombe PA, Omary MB. 'Hard' and 'soft' principles defining the structure, function and regulation of keratin intermediate filaments. Curr Opin Cell Biol. 2002; 14: 110–122. PMID: 11792552
4. Fuchs E, Green H. Changes in keratin gene expression during terminal differentiation of the keratinocyte. Cell 1980; 19: 1033–1042. PMID: 6155214
5. Homberg M, Magin TM. Beyond expectations: novel insights into epidermal keratin function and regulation. Int Rev Cell Mol Biol. 2014; 311: 265–306. https://doi.org/10.1016/B978-0-12-800179-0.00007-6 PMID: 24952920
6. Pan X, Hobbs RP, Coulombe PA. The expanding significance of keratin intermediate filaments in normal and diseased epithelia. Curr Opin Cell Biol. 2013; 25: 47–56. https://doi.org/10.1016/j.ceb.2012.10.018 PMID: 23270662
7. Paramio JM, Casanova ML, Segrelles C, Mittnacht S, Lane EB, Jorcano JL. Modulation of cell proliferation by cytokeratins K10 and K16. Mol Cell Biol. 1999; 19: 3086–3094. PMID: 10082575
8. Oshima RG. Intermediate filaments: a historical perspective. Exp Cell Res. 2007; 313: 1981–1994. https://doi.org/10.1016/j.yexcr.2007.04.007 PMID: 17493611
9. Vaidya MM, Kanojia D.Keratins: markers of cell differentiation or regulators of cell differentiation? J Biosci. 2007; 32: 629–634. PMID: 17762135
10. Gu LH, Coulombe PA. Keratin function in skin epithelia: a broadening palette with surprising shades. Curr Opin Cell Biol. 2007; 19: 13–23. https://doi.org/10.1016/j.ceb.2006.12.007 PMID: 17178453
11. Kuga T, Sasaki M, Mikami T, Miake Y, Adachi J, Shimizu M, et al. FAM83H and casein kinase I regulate the organization of the keratin cytoskeleton and formation of desmosomes. Sci Rep. 2016; 6: 26557. https://doi.org/10.1038/srep26557 PMID: 27222304
12. Bray DJ, Walsh TR, Noro MG, Notman R. Complete structure of an epithelial keratin dimer: Implications for intermediate filament assembly. PLoS One 2015; 10: e0132706. https://doi.org/10.1371/journal.pone.0132706 PMID: 26181054
13. Schweizer J, Langbein L, Rogers MA, Winter H. Hair follicle-specific keratins and their diseases. Exp Cell Res. 2007; 313: 2010–2020. https://doi.org/10.1016/j.yexcr.2007.02.032 PMID: 17428470
14. NCBI Homo sapiens Annotation Release 107. 2015. http://www.ncbi.nlm.nih.gov/genome/annotation_euk/Homo_sapiens/107/. Accessed 02 Sep 2016.
15. Schweizer J, Bowden PE, Coulombe PA, Langbein L, Lane EB, Magin TM, Maltais L, Omary MB, Parry DA, Rogers MA, Wright MW. New consensus nomenclature for mammalian keratins. J Cell Biol. 2006; 174: 169–174. https://doi.org/10.1083/jcb.200603161 PMID: 16831889
16. Moll R, Divo M, Langbein L. The human keratins: biology and pathology. Histochem Cell Biol. 2008; 129: 705–733. https://doi.org/10.1007/s00418-008-0435-6 PMID: 18461349
17. Hesse M, Zimek A, Weber K, Magin TM. Comprehensive analysis of keratin gene clusters in humans and rodents. Eur J Cell Biol. 2004; 83:19–26. https://doi.org/10.1078/0171-9335-00354 PMID: 15085952
18. Knöbel M, O'Toole EA, Smith FJ. Keratins and skin disease. Cell Tissue Res. 2015; 360: 583–589. https://doi.org/10.1007/s00441-014-2105-4 PMID: 25620412
19. Lemke JR, Kernland-Lang K, Hörtnagel K, Itin P. Monogenic human skin disorders. Dermatology. 2014; 229: 55–64. https://doi.org/10.1159/000362200 PMID: 25012694
20. Wang F, Zieman A, Coulombe PA. Skin keratins. Methods Enzymol. 2016; 568:303–50. https://doi.org/10.1016/bs.mie.2015.09.032 PMID: 26795476

21. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 2005; 438: 803–819. https://doi.org/10.1038/nature04338 PMID: 16341006

22. Hoeppner MP, Lundquist A, Pirun M, Meadows JR, Zamani N, Johnson J, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. PLoS One 2014; 9: e91172. https://doi.org/10.1371/journal.pone.0091172 PMID: 24625832

23. Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 2009; 326: 865–867. https://doi.org/10.1126/science.1178158 PMID: 19892987

24. Coleman SJ, Zeng Z, Wang K, Luo S, Khrebtukova I, Mienaltowski MJ, et al. Structural annotation of equine protein-coding genes determined by mRNA sequencing. Anim Genet. 2010; 41 Suppl 2: 121–130.

25. Avci P, Sadasivam M, Gupta A, De Melo WC, Huang YY, Yin R, et al. Animal models of skin disease for drug discovery. Expert Opin Drug Discov. 2013; 8: 331–355. https://doi.org/10.1517/17460441.2013.761202 PMID: 23293893

26. O'Leary NA, Wright MW, Brister JR, Ciufo S, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016; 44(D1): D733–745. https://doi.org/10.1093/nar/gkv1189 PMID: 26553804

27. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29: 15–21. https://doi.org/10.1093/bioinformatics/bts635 PMID: 23104886

28. Quinlan AR. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Curr Protoc Bioinformatics. 2014; 47: 11.12.1–34.

29. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25: 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 PMID: 19505943

30. NCBI Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/). Accessed at 13 Jun 2017.

31. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013; 14: 178–192. https://doi.org/10.1093/bib/bbs017 PMID: 22517427

32. Carter RA, Shekk V, de Laat MA, Pollitt CC, Galantino-Homer HL. Novel keratins identified by quantitative proteomic analysis as the major cytoskeletal proteins of equine (Equus caballus) hoof lamellar tissue. J Anim Sci. 2010; 88: 3843–3855. https://doi.org/10.2527/jas.2010-2964 PMID: 20622188

33. Reis A, Hennies HC, Langbein L, Digweed M, Mischke D, Drechsler M, et al. Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). Nat Genet. 1994; 6: 174–179. https://doi.org/10.1038/ng0294-174 PMID: 7512862

Discussion and Perspectives

During my PhD studies, I contributed to identifying the candidate causative genetic variant for 12 skin disorders or traits related to skin and its appendages in domestic animals (Table 1). Based on their molecular diagnosis, the investigated disorders fall into different groups of monogenic skin disorders and affect diverse biological processes in the skin as shown in Figure 13.

Table 1: Genes harboring candidate causative variants in investigated phenotypes and their corresponding human disorders (OMIM#).

Gene	Phenotype	Species	Human disorder (OMIM#)
ASPRV1	ichthyosis	dog	?
COL5A1	Ehlers-Danlos syndrome	cat	130000
EDA	X-linked hypohidrotic ectodermal dysplasia	dog	305100
KRT71	curly hair (and follicular dysplasia?)	dog	615896
MKLN1	lethal acrodermatitis	dog	?
MLPH	colour dilution	dog	609227
NSDHL	congenital cornification disorders	dog, cat	308050, 300831
OCA2	oculocutaneous albinism, type 2	dog	203200, 227220
SLC45A2	oculocutaneous albinism, type 4	dog	227240, 66574
ST14	naked foal syndrome	horse	602400
SUV39H2	hereditary nasal parakeratosis	dog	?

Five of the disorders can be classified as ichthyoses / generalized Mendelian disorders of cornification: The first disorder is the sporadic ichthyosis in a German Shepherd most likely caused by a *de novo* mutation event in the *ASPRV1* gene encoding an aspartic protease involved in profilaggrin-to-filaggrin processing. Second and third, the congenital cornification disorders with characteristic epidermal lesions following the Blaschko's lines in two female Labrador (-cross) dogs and a female cat fall into this category. They represent however also disorders of functional X-chromosomal mosaicism, as the identified candidate causative variants are located in the *NSDHL* gene on the X-chromosome, and early embryonic inactivation of one X chromosome with either wildtype or mutant allele leads to the characteristic mosaic pattern. Fourth, horses with naked foal syndrome, where hairlessness is the most striking feature at first glance, turned out to also show signs of ichthyosis like dry and scaly skin. All affected horses were found to be homozygous for a mutant *ST14* allele. *ST14* codes for an integral membrane serine protease critical for barrier formation in the epidermis. Finally, fifth, hereditary nasal parakeratosis is caused by variants in the *SUV39H2* gene. *SUV39H2* encodes a histone lysine methyltransferase involved in the epigenetic regulation of keratinocyte differentiation (145).

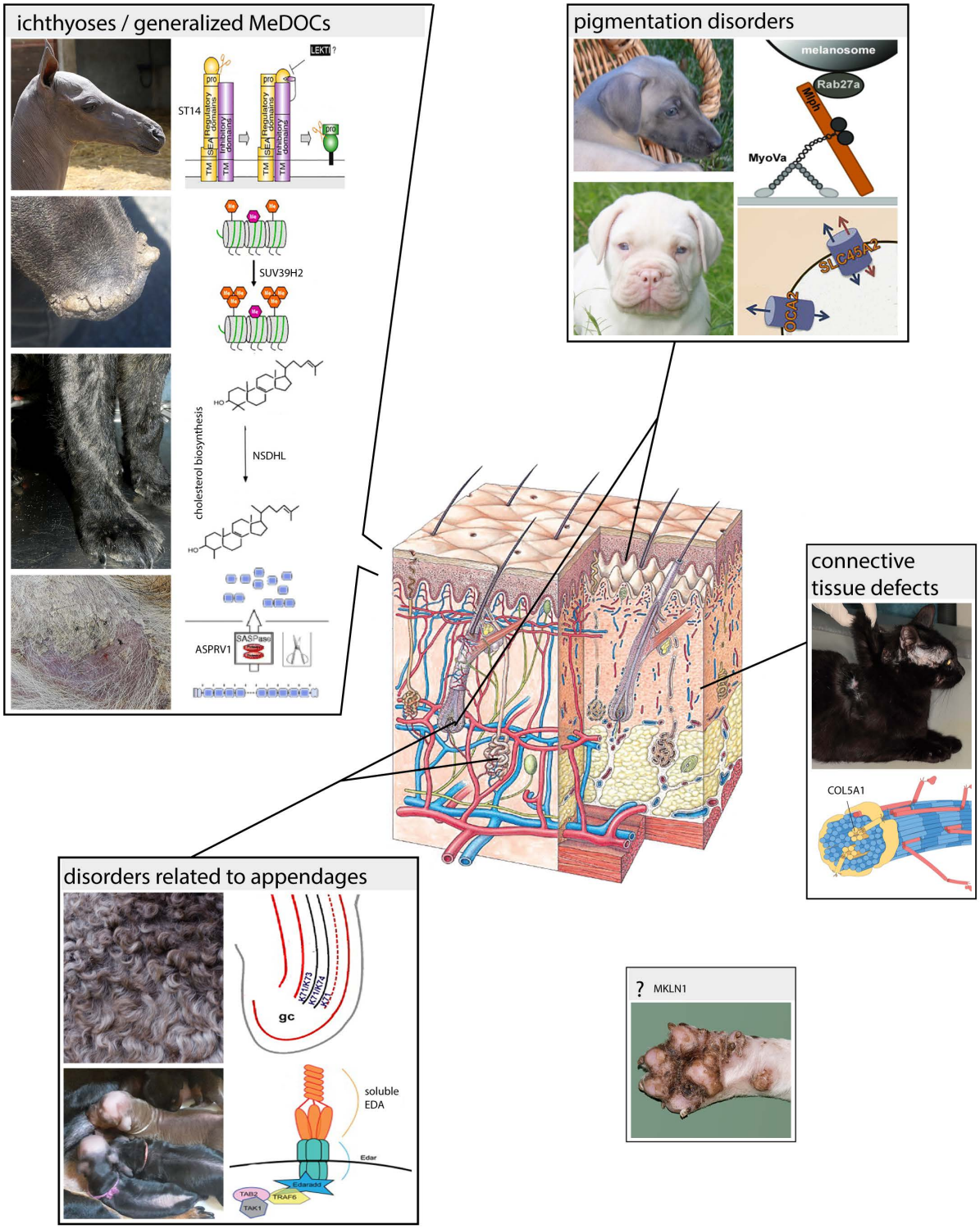


Figure 13: Investigated skin-related phenotypes and suggested underlying genetic defects. Defects in different identified genetic variants affect different skin compartments. Disorders are grouped according to the classification by Lemke and collegues (6). The classification of lethal acrodermatitis and the identified splice defect in *MKLN1* is unclear. For oculocutaneous albinism and cornification disorders due to *NSDHL* variants, only one case is shown. MeDOCs = Mendelian disorders of cornification. Pictures in this illustration are adapted with modifications from (1, 25, 91, 146-161).

One of the investigated disorders, namely Ehlers-Danlos syndrome, is a genodermatosis caused by a connective tissue defect. In the affected cat described in the results section, this defect is most likely due to an insufficient amount of collagen type V alpha I chains (haploinsufficiency).

X-linked hypohidrotic ectodermal dysplasia caused by genetic variants in the *EDA* gene and the curly hair, most likely due to a defective keratin 71 protein in dogs, are disorders/traits of ectodermal appendages. While keratin 71 is specifically expressed in the inner root sheath of the hair follicle and only seems to affect hair follicle morphology and hair texture, the broad role of ectodysplasin A during morphogenesis of ectodermal appendages is reflected by defects in multiple organs in the investigated affected dogs.

Three of the investigated traits in dogs were related to coat colour, and can therefore be classified as pigmentation disorders or more specifically, disorders with hypopigmentation. (In domestic animals, alterations in the pigmentation without effects on health are not considered disorders. Such coat color mutants are often highly valued by human breeders.) The *MLPH* variant identified in the study on Chow Chows only leads to a diluted coat colour due to an impaired melanosome transport in the skin. In the two other disorders, where defects in *SLC45A2* and *OCA2* were identified, the dogs also showed signs of photophobia besides a coat colour dilution.

Finally, the role of *MKLN1* in skin and the classification of lethal acrodermatitis are unclear. The encoded protein muskelin 1 was originally reported to be involved in responses to thrombospondin 1, an extracellular matrix component (162), which however evolved after muskelin 1 (162, 163). Since this initial characterization, diverse functions and binding partners for the ubiquitously expressed multidomain protein muskelin 1 have been discussed. The reported binding partners include EP3α prostaglandin receptor, the cyclin-dependent kinase 5 activator p39, heme-oxidase 1 and the cardiogenic T-box transcription factor TBX20 (164-167). Furthermore, by interacting with RanBP9/RanBPM as part of the CTLH complex, the yeast E3 ubiquitin ligase homolog, muskelin 1 was reported to regulate cell morphology (168, 169). More recent reports highlighted a role for muskelin 1 as transportation factor (170, 171). In mouse neurons it was demonstrated that muskelin 1 binds to the $GABA_A$ receptor during transport on both actin and microtubules, and is required for $GABA_A$ receptor endocytosis. *Mkln1*$^{-/-}$ mice showed a coat colour dilution, suggesting that muskelin 1 is expressed in different tissues and involved in trafficking of other cargoes such as melanosomes (171). This year, muskelin 1 was shown to be involved in bidirectional prion protein vesicle trafficking. In the proposed model, muskelin 1 controls the balance between intracellular lysosomal degradation and extracellular exosome trafficking and is thus involved in the regulation of prion disease progression (170). With our study on LAD in dogs, a novel function for muskelin 1 in skin was disclosed. Similarly to *Mkln1*$^{-/-}$ mice, the coat colour of affected dogs was diluted but

among the additional clinical signs and most typical for LAD, skin lesions of the paws and the face were observed, which was not reported in mice.

As indicated in Table 1, most of the identified candidate causative variants were located in genes already known to be mutated in the corresponding human disorders. However, given the different disease manifestations in domestic animals and humans, in most cases the molecular diagnosis rather than the clinical signs led to this conclusion. As an example, during the investigation of the congenital cornification disorder in a female Labrador retriever and her cross-bred daughter, the lesions following the lines of Blaschko and the lethality in male littermates indicated an X-linked disorder. The *NSDHL* gene, which turned out to harbor a large structural variant in the two affected dogs, was however not considered to be a particularly well fitting candidate gene, given that mutant *NSDHL* alleles in humans typically cause CHILD syndrome. Characteristic for this human disorder, but not present in all cases, are the limb defects and the unilateral distribution of the lesions (172, 173). Both characteristics were not present in the affected dogs. The difficulty in linking the clinical signs in affected domestic animals to known genodermatoses in humans is not surprising, given that diagnosis in humans is already challenging due to heterogeneous phenotypes in one disorder, overlapping phenotypes between different disorders and the rarity of genodermatoses (174).

In three of the studied disorders, I identified a candidate causative variant in a gene that is not known to be involved in human genetic skin disorders. One of the genes, *SUV39H2*, was already reported to be mutated in Labrador Retrievers with hereditary nasal parakeratosis (145). The identification of a second, independently occurring variant in Greyhounds affected with a very similar condition therefore adds evidence to the causal relationship between *SUV39H2* variants and the trait. The fact that in the framework of my PhD studies, as well as in other studies, new candidate genes for genodermatoses were first detected in domestic animal models, highlights the role for spontaneous domestic animal mutants as models for human skin disorders.

During the last years, purebred animals have gained attention as models for human disorders. Especially the dog has several characteristics that make it an excellent model for human genetic disorders: Dogs are clinically and physiologically similar to humans, are companion animals that share the living environment with humans who are usually willing to provide health and ancestry information, and they have a unique population structure and genomic architecture (175). Today's purebred dogs are kept in closed populations, comparable to geographically isolated human populations, but isolated more extremely. The dogs of the ~400

modern dog breeds differ greatly from each other in morphology but also personality, leading to a large variety in phenotypes (176, 177). This large diversity is however a diversity between breeds rather than within breeds. Due to founder effects, bottlenecks or popular sire effects, many of the breeds are enriched for desired alleles, but also for disease alleles. This explains why about half of the known canine genetic disorders are specific to one breed or show a breed predisposition (177, 178). Many of these genetic disorders are shared between dogs and humans (179). Compared to humans, the identification of a causative variant in dogs is however generally more straightforward for the following reasons: First, there is a large diversity of breeds with well defined characteristics, enabling studies within and across breeds. Second, if the disease is specific or more frequent in one breed, and therefore in a relatively small (effective) population, it is likely that there is only one disease alleles that is identical by descent in all affected cases. Third, the genetic diversity within a breed is low, and linkage disequilibrium in dog breeds is ~100x higher than in humans, which means that less genetic markers are required in association studies, making them a powerful tool (176, 180). Fourth, large families and the available pedigree information also facilitate family based methods.

Beside the mentioned advantages, genetic analyses in non-model or non-human organisms can be challenging due to limited availability and lower quality of resources. As an example, the genetic analyses performed during my studies strongly depend on the quality of the reference genome assembly as well as the gene annotation. The first canine high-quality draft genome was released in 2005, and in the current reference assembly, CanFam 3.1 thousands of gaps were filled and the quality further improved (181). However, as demonstrated once more during the investigation on a German Spitz family with oculocutaneous albinism, the current canine reference assembly still contains errors. Assembly and annotation errors can be solved to a large extent, this requires however human expert data interpretation. In such regions, identification of causative variants can be challenging when missed by automated bioinformatics data analyses. Similarly, large structural variants such as deletions, insertions, and especially variants due to the insertion of transposable elements such as e.g. SINEs (176, 182), are missed by calling short SNVs and indel variants with our standard automated pipeline. Therefore, in cases with sporadic disorders such as the German Shepherd with ichthyosis, where we identified a *de novo* missense variant in the *ASPRV1* gene that was not a candidate gene for ichthyosis, it would have been impossible to find the genetic cause if it was a structural variant and missed by the bioinformatics pipeline.

For all but one listed disorder, the identified candidate causative variant was likely to affect protein function and was perfectly associated with the phenotype in the study population. This is consistent with Mendelian disorders that are usually due to rare, deleterious, highly

penetrant alleles (183). The assumption that a disease allele is rare and only present in affected animals but not in unrelated controls makes whole genome sequencing alone or combined with genetic mapping an extremely powerful approach, even with a small number of cases. This assumption is however not met in complex disorders. In complex disorders, usually a large sample size is required to explain only a small proportion of a disorder's heritability by association studies. Follicular dysplasia seems to be such a complex disorder. The *KRT71* variant that we identified as fixed variant in the Curly Coated Retriever breed is likely a predisposing factor for the follicular dysplasia. To strengthen this hypothesis and to identify additional, potentially common alleles or even environmental factors' contribution to the disease, a larger, better phenotyped sample cohort is required. Furthermore, it is likely that variants contributing to complex diseases are located in non-coding regions and these would be missed by prioritizing protein-changing variants as I did for monogenic disorders investigated in this thesis.

Most of my genetic analyses ended with the identification and publication of compelling candidate variants and thus provide a valuable basis for further investigations. They do however not prove causality and give at most a superficial insight into disease pathogenesis. According to Marian and colleagues, and analogous to Koch's postulates for infectious diseases, the following four conditions should be met to prove causality: 1) The variant must be present and enriched in cases or families with the phenotype. 2) The variant must be pathogenic and functional, meaning rare or novel, affecting a conserved position and being protein-changing. 3) A model system without the phenotype should show a comparable phenotype upon introducing the variant. 4) After removing the variant, the phenotype should be reversed (184). During my PhD studies, I could partially fulfill the first two requirements with genetic data and simple experiments on RNA or protein level. Addressing the third and fourth conditions however will require more sophisticated functional experiments with tools such as CRISPR/Cas9, enabling targeted genome editing.

In conclusion, this thesis showed that a combination of whole genome sequencing and genetic mapping is a powerful approach to identify candidate causative variants for monogenic skin disorders in domestic animals. The identified variants provide the basis for genetic testing, which enables controlled breeding and eradication of monogenic disorder as a long term goal. In purebred animals, the analysis of monogenic disorders is straightforward due to unique genetic architecture, population structure and breeding practices. This allows the identification of new candidate genes for genetic disorders that have not been reported in humans before. Thus, investigating genodermatoses in animal models can lead to new insights into biological functions and new candidate genes for genodermatoses in humans and other species.

Acknowledgments

About three years ago I started my position at the Institute of Genetics and was overwhelmed with all the new methods, scientific terms, people and responsibilities. Starting in a new place was not easy - but to leave this place that has become everyday life and home at the same time will be even more difficult.

I would like to thank Prof Dr.Tosso Leeb for being the most supportive supervisor I can imagine. Thank you for giving me the freedom to pursue my own ideas while at the same time pushing me when I needed to be pushed. I am deeply impressed by your knowledge, professionalism and commitment but also by your humaneness and handling of difficult situations. Your advice has helped me in my development as scientist and human being.

Thank you to my thesis committee; Dr. Arnaud Galichet for being my co-supervisor and Prof. Dr. Carlo Largiadèr for being my mentor. I would also like to thank Prof. Dr. Carrie Finno for agreeing to act as external co-referee.

I would like to thank the whole DermFocus team for the help with my research, especially Prof. Dr. Eliane Müller, Prof. Dr. Monika Welle, Prof. Dr. Petra Roosje, and for their valuable samples, histopathological reports, pictures and great ideas.

Thank you to the lab technicians, especially Fabiana Jakob, Sabrina Schenk, Nathalie Besuchet-Schmutz and Muriel Fragnière- without you, nothing would work!

I would like to thank the bioinformaticians, Dr. Vidhja Jagannathan and Dr. Irene Keller for handling the big data and creating files that non-bioinformatics/magicians are able to understand.

Thank you to all collaborators in the different research projects. It was a pleasure and extremely interesting to work internationally together with great scientists.

My warmest thank you goes to my family and friends for always being there for me, and to all of my collegues at the Institute of Genetics. A pleasant working atmosphere is essential to me, and working with you was more than pleasant! Thank you so much!

I would also like to thank the Swiss National Science Foundation (SNSF) for funding this project.

Curriculum vitae

Removed due to data privacy reasons

List of publications

Published articles

Karli P, Oevermann A, **Bauer A**, Jagannathan V, Leeb T. *MFSD8* single-base pair deletion in a Chihuahua with neuronal ceroid lipofuscinosis. Animal Genetics. 2016;47(5):631.

Bauer A, Waluk DP, Galichet A, Timm K, Jagannathan V, Sayar BS, Wiener DJ, Dietschi E, Müller EJ, Roosje P, Welle MM, Leeb T. A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. PLoS Genetics. 2017;13(3):e1006651.

Bauer A, Hiemesch T, Jagannathan V, Neuditschko M, Bachmann I, Rieder S, Mikko S, Penedo MC, Tarasova N, Vitková M, Sirtori N, Roccabianca P, Leeb T, Welle MM. A nonsense variant in the *ST14* gene in Akhal-Teke Horses with Naked Foal Syndrome. G3 (Bethesda, Md). 2017;7(4):1315-21.

Balmer P, **Bauer A**, Pujar S, McGarvey KM, Welle M, Galichet A, Müller EJ, Pruitt KD, Leeb T, Jagannathan V. A curated catalog of canine and equine keratin genes. PloS One. 2017;12(8):e0180359.

Bauer A, De Lucia M, Jagannathan V, Mezzalira G, Casal ML, Welle MM, Leeb T. A large deletion in the *NSDHL* gene in Labrador Retrievers with a congenital cornification disorder. G3 (Bethesda, Md). 2017;7(9):3115-21.

Caduff M, **Bauer A**, Jagannathan V, Leeb T. A single base deletion in the *SLC45A2* gene in a Bullmastiff with oculocutaneous albinism. Animal Genetics. 2017;48(5):619-21.

Caduff M, **Bauer A**, Jagannathan V, Leeb T. *OCA2* splice site variant in German Spitz dogs with oculocutaneous albinism. PloS One. 2017;12(10):e0185944.

Bauer A, Kehl A, Jagannathan V, Leeb T. A novel *MLPH* variant in dogs with coat colour dilution. Animal Genetics. 2018;49(1):94-7.

Bauer A, Jagannathan V, Hogler S, Richter B, McEwan NA, Thomas A, Cadieu E, André C, Hytönen MK, Lohi H, Welle MM, Roosje P, Mellersh C, Casal ML, Leeb T. *MKLN1* splicing defect in dogs with lethal acrodermatitis. PLoS Genetics. 2018;14(3):e1007264.

Bauer A, Nimmo J, Newman R, Brunner M, Welle MM, Jagannathan V, Leeb T. A splice site variant in the *SUV39H2* gene in Greyhounds with nasal parakeratosis. Animal Genetics. 2018;49(2):137-40.

Spycher M, **Bauer A**, Jagannathan V, Frizzi M, De Lucia M, Leeb T. A frameshift variant in the *COL5A1* gene in a cat with Ehlers-Danlos syndrome. Animal Genetics. Animal Genetics. 2018;49(6):641-644

Hadji Rasouliha S, **Bauer A**, Dettwiler M, Welle MM, Leeb T. A frameshift variant in the *EDA* gene in Dachshunds with X-linked hypohidrotic ectodermal dysplasia. Animal Genetics. 2018;49(6):651-654.

Bauer A, Hadji Rasouliha S, Brunner MAT, Jagannathan V, Bucher I, Bannoehr J, Varjonen K, Bond R, Bergvall K, Welle MM, Roosje P, Leeb T. A second *KRT71* allele in curly coated dogs. Animal Genetics. doi: 10.1111/age.12743

De Lucia M, **Bauer A**, Spycher M, Jagannathan V, Romano E, Welle MM, Leeb T. Genetic variant in the *NSDHL* gene in a cat with multiple congenital lesions resembling inflammatory linear verrucous epidermal nevi (ILVEN). Veterinary Dermatology. doi: 10.1111/vde.12699

Unpublished articles

De Lucia M, Angileri M, **Bauer A**, Spycher M, Jagannathan V, Denti D, Di Diodoro F, Ferro S, Mezzalira G, Welle MM, Leeb T. X-linked cutaneous mosaicism in animals. Veterinary Dermatology. *In revision*

Brunner MAT, Rüfenacht S, **Bauer A**, Erpel S, Buchs N, Braga-Lagache S, Manfred Heller M, Leeb T, Jagannathan V, Wiener DJ, Welle MM. Bald Thigh Syndrome in sighthounds – revisiting the cause of a well-known disease. *Submitted*

Conference abstracts

Bachmann I, **Bauer A**, Welle MM, Leeb T. Haarlose Fohlen beim Achal Tekkiner. Agroscope Science. 2016;32:36-37

Bauer A, Hiemesch T, Jagannathan V, Neuditschko M, Rieder S, Bachman I, Mikko S, Penedo C, Tarasova N, Vitková M, Sirtori N, Roccabianca P, Leeb T, Welle MM. Eine Variante im *ST14* Gen führt zu erblicher Haarlosigkeit bei Achal-Tekkinern. Agroscope Science. 2017;48:16-17

Bauer A, Balmer P, Brunner MAT, Caduff M, De Lucia M, Drögemüller C, Drögemüller M, Jagannathan V, Nimmo J, Murgiano L, Sayar BS, Tarasova N, Timm K, Towers RE, Zur G, Müller E, Roosje P, Welle MM, Leeb T. Genetic analysis of genodermatoses in domestic animals. European Human Genetics Conference 2018

Appendix

Tables A1 - A9. Compilation of human and non-laboratory animal genodermatoses. The list of human disorders is based on Lemke et al. (6) with modifications. Disorders in non-laboratory animals are based on the OMIA (online Mendelian inheritance in animals) database (http://omia.org). Entries of the category «colour» and results from searches with the terms «skin» and «hair» as well as the genes listed among the human disorders were considered. Please note that this list might not comprise all known genodermatoses and that some listed phenotypes are not considered a disorder. AD = autosomal dominant, AR = autosomal recessive, ASD = autosomal semi-dominant, CNV = copy number variant, MOI = mode of inheritace, XL = X-linked, ZL = Z-linked

Table A1. Inherited ichthyoses / generalized Mendelian disorders of cornification

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
ABCA12	Ichthyosis, autosomal recessive 4B (harlequin)	242500	AR	Ichthyosis congenita	000547	Cattle	AR
	Ichthyosis, congenital, autosomal recessive 4A	601277	AR				
ABHD5	Chanarin-Dorfman syndrome	275630	AR				
ALDH3A2	Sjögren-Larsson syndrome	270200	AR				
ALOX12B	Ichthyosis, congenital, autosomal recessive 2	242100	AR				
ALOXE3	Ichthyosis, congenital, autosomal recessive 3	606545	AR				
AP1S1	MEDNIK syndrome	609313	AR				
AQP5	Palmoplantar keratoderma, Bothnian type	600231	AD				
ASPRV1				Ichthyosis	002099	Dog	AD
ATP2A2	Acrokeratosis verruciformis	101900	AD				
	Darier disease	124200	AD				
ATP2C1	Hailey-Hailey disease	169600	AD				
CDSN	Ichthyosis-hypotrichosis	146520	AD/AR				
CERS3	Autosomal recessive congenital ichthyosis	615023	AR				
CLDN1	ILVASC	607626	AR				
COL14A1	Keratoderma, palmoplantar, punctate type IB	614936	AD				
CSTA	Exfoliative ichthyosis	607936	AR				
CTSC	Papillon-Lefèvre syndrome	245000	AR				
CYP4F22	Ichthyosis, congenital, autosomal recessive 5	604777	AR				
DKC1	Dyskeratosis congenita	305000	XL				
DSG1	Striate palmoplantar keratoderma (PPKS1)	148700	AD				
DSP	Striate palmoplantar keratoderma (PPKS2)	612908	AD				
EBP	Chondrodysplasia punctata, X-linked dominant	302960	XL				
ELOVL4	Ichthyosis, spastic quadriplegia and mental retardation	614457	AR				
ENPP1	Cole disease		AD				
ERCC2	Trichothiodystrophy	601675	AR				
ERCC3	Trichothiodystrophy	601675	AR				
FLG	Ichthyosis vulgaris	146700	AD				
GJB2	KID syndrome	148210	AD				
GJB3	Erythrokeratodermia variabilis et progressiva	133200	AD				
GJB4	Erythrokeratodermia variabilis with erythema gyratum repens	133200	AD				

Table A1. Inherited ichthyoses / generalized Mendelian disorders of cornification

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
GJB6	Ectodermal dysplasia, Clouston type	129500	AD				
GTF2H5	Trichothiodystrophy	601675	AR				
JUP	Naxos disease	601214	AR				
KRT1	Ichthyosis, cyclic, with epidermolytic hyperkeratosis	607602	AD				
KRT10	Ichthyosis, cyclic, with epidermolytic hyperkeratosis	607602	AD	Hyperkeratosis, epidermolytic	001415	Dog	AR
KRT16	Pachyonychia congenita type I (Jadassohn-Lewandowsky syndrome)	148067	AD	Palmoplantar keratoderma, nonepidermolytic, focal 1	002088	Dog	AR
KRT17	Pachyonychia congenita type II (Jackson-Lawler syndrome)	167210	AD				
KRT2	Ichthyosis bullosa of Siemens	146800	AD				
KRT6A	Pachyonychia congenita type I (Jadassohn-Lewandowsky syndrome)	167200	AD				
KRT6B	Pachyonychia congenita type II (Jackson-Lawler syndrome)	167210	AD				
KRT9	Epidermolytic palmoplantar keratoderma	144200	AD				
LIPN	Ichthyosis, congenital, autosomal recessive 8	613943	AR				
LOR	Vohwinkel syndrome (ichthyotic variant)	604117	AD				
MBTPS2	IFAP syndrome with or without BRESHECK syndrome Keratosis follicularis spinulosa decalvans, X-linked	308205 308800	XL XL	Brindle 1	002021	Horse	XL
MKLN1				Lethal acrodermatitis	002146	Dog	AR
MPLKIP	Trichothiodystrophy, non-photosensitive 1	234050	AR				
NHP2	Dyskeratosis congenita, autosomal recessive 2	613987	AR				
NIPAL4	Ichthyosis, congenital, autosomal recessive 6	612281	AR	Autosomal recessive congenital ichthyosis	001980	Dog	AR
NOP10	Dyskeratosis congenita, autosomal recessive 1	224230	AR				
NSDHL	CHILD syndrome	308050	XL	Congenital cornification disorder ILVEN-like lesions	002117	Dog Cat	XL XL
PKP1	Ectodermal dysplasia/skin fragility syndrome	604536	AR	Ectodermal dysplasia/skin fragility syndrome	001864	Dog	AR
PLD4				Zinc deficiency-like syndrome	001935	Cattle	AR
PNPLA1	Ichthyosis, congenital, autosomal recessive 10	615024	AR	Ichthyosis	001588	Dog	AR
POMP	Keratosis linearis with ichthyosis congenita and sclerosing keratoderma	601952	AR				
RHBDF2	Tylosis with oesophageal cancer	148500	AD				

Table A1. Inherited ichthyoses / generalized Mendelian disorders of cornification

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
SAT1	Keratosis follicularis spinulosa decalvans	308800	XL				
SERPINB7	Palmoplantar keratosis, Nagashima type		AR				
SLC27A4	Ichthyosis prematurity syndrome	608649	AR	Ichthyosis	001973	Dog	AR
SLC39A4	Acrodermatitis enteropathica	201100	AR	Acrodermatits enteropathica / Lethal trait A46	00593	Cattle	AR
SLURP1	Mal de Meleda	248300	AR				
SMARCAD1	Adermatoglyphia	136000	AD				
SNAP29	CEDNIK syndrome	609528	AR				
SPINK5	Netherton syndrome	256500	AR				
ST14	Ichthyosis with hypotrichosis	610765	AR	Naked foal syndrome	2096	Horse	AR
STS	Ichthyosis, X-linked	308100	XL				
SUMF1	Multiple sulphatase deficiency	272200	AR				
TAT	Richner-Hanhart syndrome, tyrosinaemia, type II	276600	AR				
TERC	Dyskeratosis congenita, autosomal dominant 1	127550	AD				
TERT	Dyskeratosis congenita, autosomal recessive 4	613989	AR				
TGM1	Ichthyosis, congenital, autosomal recessive 1	242300	AR	Ichthyosis	000546	Dog	AR
TGM5	Peeling skin syndrome, acral type	609796	AR				
TINF2	Dyskeratosis congenita, autosomal dominant 3	613990	AD				
WRAP53	Dyskeratosis congenita, autosomal recessive 3	613988	AR				

Abbreviations: MEDNIK = mental retardation, enteropathy, deafness, neuropathy, ichthyosis and keratoderma; ILVASC = ichthyosis, leucocyte vacuoles, alopecia and sclerosing cholangitis; KID = keratitis, ichthyosis and deafness; IFAP = ichthyosis follicularis, alopecia and photophobia; BRESHECK = brain anomalies, retardation, ectodermal dysplasia, skeletal malformations, Hirschsprung disease, ear/eye anomalies, cleft palate/cryptorchidism and kidney dysplasia/hypoplasia; CHILD = congenital hemidysplasia with ichthyosiform erythroderma and limb defects; CEDNIK = cerebral dysgenesis, neuropathy, ichthyosis and keratoderma.

Table A2. Inherited epidermolyses and blistering disorders

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#.	Species	MOI
COL17A1	Epidermolysis bullosa, junctional, non-Herlitz type	226650	AR				
COL7A1	Epidermolysis bullosa dystrophica, autosomal recessive	226600	AR	Epidermolysis bullosa, dystrophic	000341	Cattle,dog	AR
	Epidermolysis bullosa dystrophica, Bart type	132000	AD				
	Epidermolysis bullosa dystrophica, localisata variant		AD				
	Epidermolysis bullosa dystrophica, autosomal dominant	131750	AD				
	Epidermolysis bullosa dystrophica, autosomal recessive	226600	AR				
	Epidermolysis bullosa pruriginosa	604129	AD/AR				
	Epidermolysis bullosa, pretibial	131850	AD				
	Toenail dystrophy, isolated	607523	AD				
	Transient bullous epidermolysis of the newborn	131705	AD/AR				
DSP	Epidermolysis bullosa, lethal acantholytic	609638	AR				
EXPH5	Epidermolysis bullosa, non-specific, autosomal recessive	615028	AR				
FERMT1	Kindler syndrome	173650	AR				
ITGA3	Interstitial lung disease, nephritic syndrome and epidermolysis bullosa, congenital	614748	AR				
ITGA6	Epidermolysis bullosa, junctional, with pyloric stenosis	226730	AR				
ITGB4	Epidermolysis bullosa of hands and feet	131800	AR	Epidermolysis bullosa, junctionalis	001948	Cattle, sheep	AR
	Epidermolysis bullosa, junctional, non-Herlitz type	226650	AR				
	Epidermolysis bullosa, junctional, with pyloric atresia	226730	AR				
KRT1	Ichthyosis, cyclic, with epidermolytic hyperkeratosis	607602	AD				
KRT10	Ichthyosis, cyclic, with epidermolytic hyperkeratosis	607602	AD				
KRT5	Dowling-Degos disease 1	179850	AD	Epidermolysis bullosa, simplex	002081	Cattle	AD
	Epidermolysis bullosa simplex with migratory circinate erythema	609352	AD				
	Epidermolysis bullosa simplex with mottled pigmentation	131960	AD				
	Epidermolysis bullosa simplex, Dowling-Meara type	131760	AD				
	Epidermolysis bullosa simplex, Koebner type	131900	AD				
	Epidermolysis bullosa simplex, Weber-Cockayne type	131800	AD				
KRT14	Dermatopathia pigmentosa reticularis	125595	AD				
	Epidermolysis bullosa simplex, Dowling-Meara type	131760	AD				
	Epidermolysis bullosa simplex, Koebner type	131900	AD				
	Epidermolysis bullosa simplex, recessive 1	601001	AR				
	Epidermolysis bullosa simplex, Weber-Cockayne type	131800	AD				

Table A2. Inherited epidermolyses and blistering disorders

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#.	Species	MOI
	Naegeli-Franceschetti-Jadassohn syndrome	161000	AD				
LAMA3	Epidermolysis bullosa, generalized atrophic benign	226650	AR	Epidermolysis bullosa, junctionalis	001677	Cattle, dog, horse	AR
	Epidermolysis bullosa, junctional, Herlitz type	226700	AR				
	Laryngo-onychocutaneous syndrome	245660	AR				
LAMB3	Epidermolysis bullosa, junctional, Herlitz type	226700	AR				
	Epidermolysis bullosa, junctional, non-Herlitz type	226650	AR				
LAMC2	Epidermolysis bullosa, junctional, Herlitz type	226700	AR	Epidermolysis bullosa, junctionalis	001678	Cattle, horse, sheep	AR
	Epidermolysis bullosa, junctional, non-Herlitz type	226650	AR				
PLEC1	Epidermolysis bullosa simplex with pyloric atresia	612138	AR	Epidermolysis bullosa, simplex	002080	Dog	AR
	Epidermolysis bullosa simplex, Ogna type	131950	AD				
	Muscular dystrophy with epidermolysis bullosa simplex	226670	AR				
POFUT1	Dowling-Degos disease 1	615327	AD				
POGLUT1	Dowling-Degos disease		AD				

Table A3. Nucleotide excision repair disorders

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
DDB2	Xeroderma pigmentosum, group E, DDB-negative subtype	278740	AR	Ocular squamous cell carcinoma	000735	Horse	AR
ERCC1	Xeroderma pigmentosum		AR				
ERCC2	Xeroderma pigmentosum, group D	278730	AR				
	Trichothiodystrophy	601675	AR				
ERCC3	Xeroderma pigmentosum, group B	610651	AR				
	Trichothiodystrophy	601675	AR				
ERCC4	Xeroderma pigmentosum, group F	278760	AR				
ERCC5	Xeroderma pigmentosum, group G	278780	AR				
ERCC6	Cockayne syndrome, type B	133540	AR				
	UV-sensitive syndrome 1	600630	AR				
ERCC8	Cockayne syndrome, type A	216400	AR				
	UV-sensitive syndrome 2	614621	AR				
GTF2H5	Trichothiodystrophy	601675	AR				
MPLKIP	Trichothiodystrophy	234050	AR				
POLH	Xeroderma pigmentosum, variant type	278750	AR				
RECQL4	Rothmund-Thompson syndrome	268400	AR				
WRN	Werner syndrome	277700	AR				
XPA	Xeroderma pigmentosum, group A	278700	AR				
XPC	Xeroderma pigmentosum, group C	278720	AR				

Table A4. Inherited disorders with hypo- hyperpigmentation (Coat/feather colour in animals)

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
ADAM10	Reticulate acropigmentation of Kitamura		AD				
ADAR	Dyschromatosis symmetrica hereditaria	127400	AD				
AP3B1	Hermansky-Pudlak syndrome 2	608233	AR	Neutropenia, cyclic (Grey Collie syndrome)	000248	Dog	AR
BCO2				Skin/shank colour yellow	001449	Chicken	AR
BLM	Bloom syndrome	210900	AR				
BLOC1S3	Hermansky-Pudlak syndrome 8	614077	AR				
BLOC1S6	Hermansky-Pudlak syndrome 9	614171	AR				
BRAF	Cardiofaciocutaneous syndrome	115150	AD				
C10ORF11/ LRMDA	Albinism, oculocutaneous, type VII	615579	AR				
CBD103				Dominant black	001416	Coyote, Dog, Wolf	AD
CDKN2A				Barring, sex linked	000102	Chicken	ZL
COPA				Dominant red	001529	Cattle	AD
CORIN				Golden	002159	Tiger	AR
CYP2J19				Feather colour red	002019	Chicken	?
DTNBP1	Hermansky-Pudlak syndrome 7	614076	AR				
EDN3	Waardenburg syndrome, type 4B	613265	AR	Hyperpigmentation, Silky/Silkie (Fibromelanosis)	001671	Chicken	AD
EDNRA				Coat colour, white spotting	000214	Goat	ASD
EDNRB	Waardenburg syndrome, type 4A	277580	AR	Waardenburg syndrome, type 4A	001765	Sheep	AR
				Overo lethal white foal syndrome	000629	Horse	AR
				Overo coat colour	000629	Horse	AD
EDNRB2				Feather colour, mottling	001904	Chicken	AR
ENPP1	Cole disease		AD				
HPS1	Hermansky-Pudlak syndrome 1	203300	AR				
HPS3	Hermansky-Pudlak syndrome 3	614072	AR				
HPS4	Hermansky-Pudlak syndrome 4	614073	AR				
HPS5	Hermansky-Pudlak syndrome 5	614074	AR	oculocutaneous albinism	002116	Stickleback	AR
HPS6	Hermansky-Pudlak syndrome 6	614075	AR				
KIT	Piebaldism	172800	AD	Diverse coat colour related phenotypes		Diverse	
KITLG				Roan	001216	Cattle	ASD

Table A4. Inherited disorders with hypo- hyperpigmentation (Coat/Feather colour in animals)

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
LVRN				Tabby		Cat	AD/ AR
MAP2K1	Cardiofaciocutaneous syndrome	115150	AD				
MAP2K2	Cardiofaciocutaneous syndrome	115150	AD				
MC1R	Oculocutaneous albinism, type II, modifier of	203200	AD	Diverse coat colour related phenotypes		Diverse	
MITF	Tietz albinism-deafness syndrome	103500	AD	Coat colour, white spotting		Buffalo, cattle, dog, horse	ASD ?
	Waardenburg syndrome, type 2A	193510	AD	Depigmentation associated with microphthalmia		Cattle	AD
				Dominant white with bilateral deafness		Cattle	AD
				Hyperpigmentation, MITF-related	002129	Chicken	?
				Waardenburg syndrome, type 2A		Golden hamster, pig	AD/ AR
				Feather colour, silver		Japanese quail	ASD
				Osteopetrosis		Japanese quail	AR
MLH1	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
MLPH	Griscelli syndrome, type 3	609227	AR	Coat/feather colour, dilute	000031, 001445	Cat, cattle, chicken, dog, Japanese quail, American mink, rabbit,	AR
MSH2	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
MSH6	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
MYO5A	Griscelli syndrome, type 1	214450	AR	Lavender foal syndrome	001501	Horse	AR
NF1	Neurofibromatosis, type 1	162200	AD				
NF2	Neurofibromatosis, type 2	101000	AD				
OCA2	Albinism, oculocutaneous, type II	203200	AR	Oculocutaneous albinism	002130	Corn snake, dog, Mexican tetra	AR
	Albinism, brown oculocutaneous	203200					
PAX3	Waardenburg syndrome, type 1	193500	AD	Coat colour, white spotting, splashed white	001688	Horse	ASD
	Waardenburg syndrome, type 3	148820	AD				
PMEL				Feather colour white	000373	Chicken	AD

Table A4. Inherited disorders with hypo- hyperpigmentation (Coat/Feather colour in animals)

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
				Merle	000211	Dog	AD
PMS3	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
PRKAR1A	Carney complex type I	160980	AD				
PSMB7				Harlequin	001454	Dog	AD
PTPN11	Leopard syndrome	151100	AD				
RAB27A	Griscelli syndrome, type 2	607624	AR				
RAB38				Oculocutaneous hypopigmentation	002101	Cattle	AR
RALY				black-and-tan/saddle tan	001806	Dog	AR
RECQL4	Rothmund-Thompson syndrome	268400	AR				
SLC24A5	Oculocutaneous albinism, type VI	113750	AR	Oculocutaneous albinism, type VI	002124	Horse	AR
SLC36A1				Champagne	001263	Horse	AD
SLC45A2	Oculocutaneous albinism, type IV	606574	AR	Oculocutaneous albinism, type IV	001821	Cattle, dog, gorilla, medaka	AR
				Coat colour, cream dilution	001344	Horse	ASD
				Coat colour, white	000213	Bengal tiger	AR
				Feather colour, albinism, sex-linked, imperfect	000370	Japanese quail	ZL
				Feather colour, silver	000915	Chicken	ZL
SNAI2	Waardenburg syndrome, type 2D	608890	AR				
	Piebaldism	172800	AR				
SOX10	Waardenburg syndrome, type 4C	613266	AR	Feather colour, dark brown	001569	Chicken	AR
SPRED1	Legius syndrome	611431	AD				
STK11	Peutz-Jeghers syndrome	175200	AD				
STX17				Grey	001356	Horse	AD
TBX3				Dun	001972	Horse	AD
TRPM1				Leopard complex spotting			AD
				Night blindness		Horse	AR
TWIST2				Coat colour, white belt	001469	Cattle	AD
TYR	Albinism, oculocutaneous, type IA	203100	AR	Diverse coat colour related phenotypes		Diverse	
	Albinism, oculocutaneous, type IB	606952	AR				
TYRP1	Albinism, oculocutaneous, type III	203290	AR	Coat colour, blonde	001362	Pig	AD
				Coat colour, brown	001249	Diverse	AD/ AR
				Feather colour, light brown	002061	Sacer falcon	ZL
				Feather colour, roux	001322	Japanese quail	ZL

Abbreviations: MMR = mismatch repair

Table A5. Disorders of ectodermal appendages including ectodermal dysplasia

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
ABCC9	Cantú syndrome	239850	AD				
ALMS1	Alström syndrome	203800	AR				
APCDD1	Hypotrichosis simplex	605389	AD				
ARHGAP31	Adams-Oliver syndrome	100300	AD				
AXIN2	Ectodermal dysplasia and neoplastic syndrome	608615	AD				
BANF1	Nestor-Guillermo progeria syndrome	614008	AR				
BCS1L	Bjørnstad syndrome	262000	AR				
CDH3	Hypotrichosis with juvenile macular dystrophy	601553	AR				
CYP26C1	Focal facial dermal dysplasia 4	614974	AR				
DDX59	Orofaciodigital syndrome V	174300	AR				
DLX3	Trichodento-osseous syndrome	190320	AD	Tricho-dento-osseous-like syndrome	002109	Cattle	AD
DOCK6	Adams-Oliver syndrome	614219	AR				
DSC3	Hypotrichosis and recurrent skin vesicles	613102	AR				
DSG4	Hypotrichosis	607903	AR				
DSP	Skin fragility-woolly hair syndrome	607655	AR				
EDA	X-linked hypohidrotic ectodermal dysplasia (ED1)	305100	XL	Anhidrotic ectodermal dysplasia, X-linked hypohidrotic ectodermal dysplasia	000543	Cattle, dog	XL
EDAR	Ectodermal dysplasia, hypohidrotic	224900	AD	Anhidrotic ectodermal dysplasia	002128	Cattle	AR
				Reduced scale-3	001695	Japanese medaka	AR
EDARADD	Ectodermal dysplasia, hypohidrotic	614941	AD				
EFNB1	Craniofrontonasal dysplasia	304110	XL				
ERCC2	Trichothiodystrophy	601675	AR				
ERCC3	Trichothiodystrophy	601675	AR				
FGF5				Long hair	000439	Diverse	AD/AR
FGF20				Scaleless	000889	Chicken	AR
FOXI3				Ectodermal dysplasia	000323	Dog	ASD
FOXN1	T-cell immunodeficiency, congenital alopecia and nail dystrophy	601705	AR	Hypotrichosis, with short life expectancy	001949	Cat	AR
GDF7				Naked neck	000701	Chicken	AD
GJB6	Ectodermal dysplasia 2, Clouston type	129500	AR				
HEPHL1				Hypotrichosis	000540	Cattle	AR
HOXC13	Ectodermal dysplasia 9, hair/nail type	614931	AR	Ectodermal dysplasia 9	002157	Pg, rabbit (GMOs)	AR

Table A5. Disorders of ectodermal appendages including ectodermal dysplasia

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
HR	Atrichia with 2 ocular lesions, also called congenital alopecia universalis	209500	AR	Atrichia with papular lesions	001348	Rhesus monkey	AR
HVEC	Cleft lip/palate-ectodermal dysplasia syndrome	225060	AR				
IFT122	Cranio-ectodermal dysplasia 1	218330	AR				
IFT43	Sensenbrenner syndrome	614099	AR				
IKBKG	Anhidrotic ectodermal dysplasia with immune deficiency	300291	XL	Incontinentia pigmenti	001899	Horse	XL
IRF2BP2				woolly hair	001528	Sheep	AR
JUP	Naxos disease	601214	AR				
KCTD1	Scalp-ear-nipple syndrome	181270	AD				
KRT17	Steatocystoma multiplex Pachyonychia congenita, Jackson-Lawler type	184500 167210	AD				
KRT25				Curly coat	000245	Horse	AD
KRT27				Curly hair, karakul-type	000246	Cattle	AD
KRT71				Curly coat	000245	Dog	ASD?
				Curly coat, Devon rex	001581	Cat	AR
				Curly coat, Selkirk rex	001712	Cat	AD
				Hypotrichosis	002114	Cattle	?
				Hypotrichosis,with whiskers short and curled	001583	Cat	AR
KRT74	Woolly hair	194300	AD				
KRT75	Pseudofolliculitis barbae	612318	AD	Frizzle	000394	Chicken	AD
KRT81	Monilethrix	158000	AD				
KRT83	Monilethrix	158000	AD				
KRT85	Ectodermal dysplasia 4, hair/nail type	602032	AR				
KRT86	Monilethrix	158000	AD				
LIPH	Woolly hair	604379	AR	Rex coat	001566	Rabbit	AR
LMNA	Hutchinson-Gilford progeria	176670	AD				
LMX1B	Nail-patella syndrome	161200	AD				
LPAR6	Woolly hair Hypotrichosis 8	278150 278150	AR	Curly/woolly coat	001684	Cat	AR
MSX1	Ectodermal dysplasia 3, Witkop type	189500	AD	Limbless	000602	Chicken	AR
PDSS2				Silky/Silkie feathering	000913	Chicken	AR
PKP1	Ectodermal dysplasia/skin fragility syndrome	604536	AR	Ectodermal dysplasia/skin fragility syndrome	001864	Dog	AR
POC1A	Short stature, onychodysplasia, facial dysmorphism and hypotrichosis syndrome	614813	AR				

Table A5. Disorders of ectodermal appendages including ectodermal dysplasia

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
PORCN	Focal dermal hypoplasia	305600	XL				
PPP1R13L				Cardiomyopathy and wolly haircoat	000161	Cattle	AR
PRL				Hairy	000441	Cattle	AD
PRLR				Slick hair	001372	Cattle	AD
PVRL1	Margarita Island type of ectodermal dysplasia (ED4)	225060	AR				
RBM28	Alopecia, neurological defects and endocrinopathy syndrome	612079	AR				
RBPJ	Adams-Oliver syndrome	614814	AD				
RMRP	Cartilage hair hypoplasia	250250	AR				
RSPO2				Furnishings	001531	Dog	AD
SGK3				Recessive hypotrichosis	001279	Dog	AR
SNRPE	Hypotrichosis	615059	AD				
SOX18	Hypotrichosis-lymphoedema-telangiectasia syndrome	607823	AD				
TP63	Ankyloblepharon-ectodermal defects-cleft lip/palate syndrome	106260	AD				
TRPS1	Trichorhinophalangeal syndrome, type I	190350	AD				
TSR2				Streaked hairlessness	000542	Cattle	XL
TWIST2	Focal facial dermal dysplasia 3, Setleis type	227260	AR				
WDR35	Cranio-ectodermal dysplasia 2	613610	AR				

Table A6. Vascular disorders

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
ACVRL1	Telangiectasia, hereditary haemorrhagic, type 2	600376	AD				
C1NH	Angio-oedema, hereditary, types I and II	106100	AD				
C7ORF22	Cerebral cavernous malformations	603284	AD				
ENG	Telangiectasia, hereditary haemorrhagic, type 1	187300	AD				
FLT4	Hereditary lymphoedema type I	153100	AD				
FOXC2	Lymphoedema-distichiasis syndrome	153400	AD				
GDF2	Telangiectasia, hereditary haemorrhagic, type 5	615506	AD				
GLMN	Glomuvenous malformations (glomangiomas)	138000	AD				
GNAQ	Sturge-Weber syndrome	185300	mosaic				
	Capillary malformations, congenital, 1, somatic, mosaic	163000	mosaic				
KRIT1	Cerebral cavernous malformations	116860	AD				
PDCD10	Cerebral cavernous malformations	603285	AD				
RASA1	Capillary malformation-arteriovenous malformation	608354	AD				
	Parkes Weber syndrome	608355	AD				
TEK	Venous malformations, multiple cutaneous and mucosal	600195	AD				

Table A7. Connective tissue defects

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
ABCC6	Pseudoxanthoma elasticum	264800	AR				
ADAMTS2	Ehlers-Danlos syndrome, type VIIC	225410	AR	Ehlers-Danlos syndrome, type VII (Dermatosparaxis)	000328	Cattle, sheep	AR
ADAMTSL2	Geleophysic dysplasia 1	231050	AR	Musladin-Lueke syndrome	001509	Dog	AR
ALDH18A1	Cutis laxa	219150	AR				
ATP6V0A2	Cutis laxa	219200	AR				
ATP7A	Menkes disease	309400	XL	Menkes disease	000640	Dog	XL
B3GALT6	Ehlers-Danlos syndrome, progeroid type 2	615349	AR				
B4GALT7	Ehlers-Danlos syndrome, progeroid type 1	130070	AD	Dwarfism, Friesian	002068	Horse	AR
COL1A1	Ehlers-Danlos syndrome, type I	130000	AD	Osteogenesis imperfecta, type II	002127	Cattle	AD
	Ehlers-Danlos syndrome, type VIIA	130060	AD	Osteogenesis imperfecta, type III	002126	Dog	AD
COL1A2	Ehlers-Danlos syndrome, type VIIB	130060	AD	Osteogenesis imperfecta	002112	Dog	AD
COL3A1	Ehlers-Danlos syndrome, type IV	130050	AD				
COL5A1	Ehlers-Danlos syndrome, type II	130010	AD				
	Ehlers-Danlos syndrome, type I	130000	AD				
COL5A2	Ehlers-Danlos syndrome, type I	130000	AD				
EFEMP2	Cutis laxa	614437	AR				
ELN	Cutis laxa	123700	AD				
FBLN5	Cutis laxa	614434	AD/AR				
FBN1	Marfan syndrome	154700	AD	Marfan syndrome	000628	Cattle	AD
FBN2	Contractural arachnodactyly, congenital	121050	AD				
HAS2				Periodic Fever Syndrome	001561	Dog	CNV
LTBP4	Cutis laxa	613177	AR				
PLOD1	Ehlers-Danlos syndrome, type VI	225400	AR	Ehlers-Danlos syndrome, type VI	001982	Horse	AR
PYCR1	Cutis laxa	614438	AR				
SMAD3	Loeys-Dietz syndrome	613795	AD				
TGFB2	Loeys-Dietz syndrome	608967	AD				
TGFBR1	Loeys-Dietz syndrome	609192	AD				
TGFBR2	Loeys-Dietz syndrome	608967	AD				
TNXB	Ehlers-Danlos syndrome, autosomal dominant, hypermobility type	130020	AD				
	Ehlers-Danlos syndrome, autosomal recessive, due to tenascin-X deficiency	606408	AD				
ZMPSTE24	Restrictive dermopathy, lethal	275210	AR				

Table A8. Dermal mosaics

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
AKT1	Proteus syndrome	176920	mosaic				
BRAF	Cardiofaciocutaneous syndrome	115150	mosaic				
FGFR3	Naevus, epidermal, somatic	162900	mosaic	Chondrodysplasia, Spider lamb	001703	Sheep	AR
GNAQ	Sturge-Weber syndrome	185300	mosaic				
	Capillary malformations, congenital, 1, somatic, mosaic	163000	mosaic				
GNAS	McCune-Albright syndrome	174800	mosaic				
HRAS	Naevus, epidermal, somatic	162900	mosaic				
	Schimmelpenning-Feuerstein-Mims syndrome, somatic mosaic	163200					
IDH1	Ollier disease/Maffucci syndrome		mosaic				
IDH2	Ollier disease/Maffucci syndrome		mosaic				
KRAS	Schimmelpenning-Feuerstein-Mims syndrome, somatic mosaic	163200	mosaic				
MAP2K1	Cardiofaciocutaneous syndrome	615279	mosaic				
MAP2K2	Cardiofaciocutaneous syndrome	615280	mosaic				
NRAS	Naevus, epidermal, somatic	162900	mosaic				
PIK3CA	Naevus, epidermal, somatic	162900	mosaic				
PORCN	Focal dermal hypoplasia	305600	mosaic				
PTEN	Linear PTEN naevus	158350	mosaic				
PTPN11	Leopard syndrome	151100	mosaic				
SOX10	Giant melanocytic naevus		mosaic				
SPRED1	Legius syndrome	611431	mosaic				
TSC1	Tuberous sclerosis complex	191100	mosaic				
TSC2	Tuberous sclerosis complex	613254	mosaic				

Table A9. Genodermatoses with tumor predisposition

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
AKT1	Cowden syndrome	615109	AD				
APC	Gardner syndrome	175100	AD	Familial adenomatous polyposis	001916	Pig (GMO)	
ATM	Ataxia-telangiectasia	208900	AR	Ataxia telangiectasia	002044	Pig (GMO)	
AXIN2	Ectodermal dysplasia and neoplastic syndrome	608615	AD				
BLM	Bloom syndrome	210900	AR				
CYLD	Tricho-epithelioma, multiple familial	601606	AD				
DDB2	Xeroderma pigmentosum, group E, DDB-negative subtype	278740	AR	Ocular squamous cell carcinoma	000735	Horse	AR
DKC1	Dyskeratosis congenita	305000	XR				
ERCC2	Xeroderma pigmentosum, group D	278730	AR				
ERCC3	Xeroderma pigmentosum, group B	610651	AR				
ERCC4	Xeroderma pigmentosum, group F	278760	AR				
ERCC5	Xeroderma pigmentosum, group G	278780	AR				
FERMT1	Kindler syndrome	173650	AR				
FH	Leiomyomatosis with or without renal cell cancer	150800	AD				
FLCN	Birt-Hogg-Dubé syndrome	135150	AD	Renal cystadenocarcinoma and nodular dermatofibrosis	001335	Dog	AD
GTF2H5	Trichothiodystrophy	601675	AR				
MLH1	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
MPLKIP	Trichothiodystrophy	234050	AR				
MSH2	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
MSH6	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
NF1	Neurofibromatosis, type 1	162200	AD				
NF2	Neurofibromatosis, type 2	101000	AD				
NHP2	Dyskeratosis congenita, autosomal recessive 2	613987	AR				
NOP10	Dyskeratosis congenita, autosomal recessive 1	224230	AR				
NOTCH3	Myofibromatosis, infantile, 2	615293	AD				
PDGFRB	Myofibromatosis, infantile, 1	228550	AD				
PIK3CA	Cowden syndrome	615108	AD				
PMS2	MMR deficiency syndrome (Turcot syndrome)	276300	AR				
POLH	Xeroderma pigmentosum, variant type	278750	AR				
PRKAR1A	Carney complex type I	160980	AD				
PTCH1	Naevoid basal cell carcinoma syndrome (Gorlin syndrome)	109400	AD				
PTCH2	Familial basal cell carcinoma	605462	AD				

Table A9. Genodermatoses with tumor predisposition

Gene	Disease	Phenotype OMIM#	MOI	Animal phenotype	Phenotype OMIA#	Species	MOI
PTEN	Cowden syndrome	158350	AD	Colorectal hamartomatous polyposis and ganglioneuromatosis	001515	Dog	?
RECQL4	Rothmund-Thompson syndrome	268400	AR				
RHBDF2	Tylosis with oesophageal cancer	148500	AD				
STK11	Peutz-Jeghers syndrome	175200	AD				
TERC	Dyskeratosis congenita, autosomal dominant 1	127550	AD				
TERT	Dyskeratosis congenita, autosomal recessive 4	613989	AR				
TINF2	Dyskeratosis congenita, autosomal dominant 3	613990	AD				
TSC1	Tuberous sclerosis complex	191100	AD				
TSC2	Tuberous sclerosis complex	613254	AD				
WRAP53	Dyskeratosis congenita, autosomal recessive 3	613988	AR				
WRN	Werner syndrome	277700	AR				
XPA	Xeroderma pigmentosum, group A	278700	AR				
XPC	Xeroderma pigmentosum, group C	278720	AR				

Abbreviations: GMO = Genetically modified organism, MMR = mismatch repair

References

1. Shimizu H. Shimizu's Dermatology. 2nd ed. Chichester, West Sussex ; Hoboken, NJ: John Wiley and Sons Ltd; 2017.

2. Eckhart L, Lippens S, Tschachler E, Declercq W. Cell death by cornification. Biochimica et biophysica acta. 2013;1833(12):3471-80.

3. Pispa J, Thesleff I. Mechanisms of ectodermal organogenesis. Developmental biology. 2003;262(2):195-205.

4. Mangelsdorf S, Vergou T, Sterry W, Lademann J, Patzelt A. Comparative study of hair follicle morphology in eight mammalian species and humans. Skin research and technology : official journal of International Society for Bioengineering and the Skin (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society for Skin Imaging (ISSI). 2014;20(2):147-54.

5. Feramisco JD, Sadreyev RI, Murray ML, Grishin NV, Tsao H. Phenotypic and genotypic analyses of genetic skin disease through the Online Mendelian Inheritance in Man (OMIM) database. The Journal of investigative dermatology. 2009;129(11):2628-36.

6. Lemke JR, Kernland-Lang K, Hortnagel K, Itin P. Monogenic human skin disorders. Dermatology (Basel, Switzerland). 2014;229(2):55-64.

7. Candi E, Schmidt R, Melino G. The cornified envelope: a model of cell death in the skin. Nature reviews Molecular cell biology. 2005;6(4):328-40.

8. Madison KC. Barrier function of the skin: "la raison d'etre" of the epidermis. The Journal of investigative dermatology. 2003;121(2):231-41.

9. Schmuth M, Martinz V, Janecke AR, Fauth C, Schossig A, Zschocke J, et al. Inherited ichthyoses/generalized Mendelian disorders of cornification. European journal of human genetics : EJHG. 2013;21(2):123-33.

10. Oji V, Tadini G, Akiyama M, Blanchet Bardon C, Bodemer C, Bourrat E, et al. Revised nomenclature and classification of inherited ichthyoses: results of the First Ichthyosis Consensus Conference in Soreze 2009. Journal of the American Academy of Dermatology. 2010;63(4):607-41.

11. Yoneda K. Inherited ichthyosis: Syndromic forms. The Journal of dermatology. 2016;43(3):252-63.

12. Takeichi T, Akiyama M. Inherited ichthyosis: Non-syndromic forms. The Journal of dermatology. 2016;43(3):242-51.

13. Matsui T, Amagai M. Dissecting the formation, structure and barrier function of the stratum corneum. International immunology. 2015;27(6):269-80.

14. Akiyama M, Shimizu H. An update on molecular aspects of the non-syndromic ichthyoses. Experimental dermatology. 2008;17(5):373-82.

15. Ishida-Yamamoto A, Igawa S. The biology and regulation of corneodesmosomes. Cell and tissue research. 2015;360(3):477-82.

16. Vahlquist A, Fischer J, Torma H. Inherited Nonsyndromic Ichthyoses: An Update on Pathophysiology, Diagnosis and Treatment. 2018;19(1):51-66.

17. Knapp AC, Franke WW, Heid H, Hatzfeld M, Jorcano JL, Moll R. Cytokeratin No. 9, an epidermal type I keratin characteristic of a special program of keratinocyte differentiation displaying body site specificity. The Journal of cell biology. 1986;103(2):657-67.

18. Swensson O, Langbein L, McMillan JR, Stevens HP, Leigh IM, McLean WH, et al. Specialized keratin expression pattern in human ridged skin as an adaptation to high physical stress. The British journal of dermatology. 1998;139(5):767-75.

19. Reis A, Hennies HC, Langbein L, Digweed M, Mischke D, Drechsler M, et al. Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK). Nature genetics. 1994;6(2):174-9.

20. Tsubota A, Akiyama M, Kanitakis J, Sakai K, Nomura T, Claudy A, et al. Mild recessive bullous congenital ichthyosiform erythroderma due to a previously unidentified homozygous keratin 10 nonsense mutation. The Journal of investigative dermatology. 2008;128(7):1648-52.

21. Terheyden P, Grimberg G, Hausser I, Rose C, Korge BP, Krieg T, et al. Recessive epidermolytic hyperkeratosis caused by a previously unreported termination codon mutation in the keratin 10 gene. The Journal of investigative dermatology. 2009;129(11):2721-3.

22. Rothnagel JA, Dominey AM, Dempsey LD, Longley MA, Greenhalgh DA, Gagne TA, et al. Mutations in the rod domains of keratins 1 and 10 in epidermolytic hyperkeratosis. Science (New York, NY). 1992;257(5073):1128-30.

23. Muller FB, Huber M, Kinaciyan T, Hausser I, Schaffrath C, Krieg T, et al. A human keratin 10 knockout causes recessive epidermolytic hyperkeratosis. Human molecular genetics. 2006;15(7):1133-41.

24. Gutierrez JA, Hannoush ZC, Vargas LG, Momany A, Garcia CC, Murray JC, et al. A Novel non-sense Mutation in Keratin 10 Causes a Familial Case of Recessive Epidermolytic Ichthyosis. Molecular genetics & genomic medicine. 2013;1(2):108-12.

25. Matsui T, Miyamoto K, Kubo A, Kawasaki H, Ebihara T, Hata K, et al. SASPase regulates stratum corneum hydration through profilaggrin-to-filaggrin processing. EMBO molecular medicine. 2011;3(6):320-33.

26. Rawlings AV, Harding CR. Moisturization and skin barrier function. Dermatologic therapy. 2004;17 Suppl 1:43-8.

27. Homberg M, Magin TM. Beyond expectations: novel insights into epidermal keratin function and regulation. International review of cell and molecular biology. 2014;311:265-306.

28. Smith FJ, Irvine AD, Terron-Kwiatkowski A, Sandilands A, Campbell LE, Zhao Y, et al. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. Nature genetics. 2006;38(3):337-42.

29. Denecker G, Hoste E, Gilbert B, Hochepied T, Ovaere P, Lippens S, et al. Caspase-14 protects against epidermal UVB photodamage and water loss. Nature cell biology. 2007;9(6):666-74.

30. Steinert PM, Marekov LN. Initiation of assembly of the cell envelope barrier structure of stratified squamous epithelia. Molecular biology of the cell. 1999;10(12):4247-61.

31. Kalinin A, Marekov LN, Steinert PM. Assembly of the epidermal cornified cell envelope. Journal of cell science. 2001;114(Pt 17):3069-70.

32. Steinert PM, Marekov LN. The proteins elafin, filaggrin, keratin intermediate filaments, loricrin, and small proline-rich proteins 1 and 2 are isodipeptide cross-linked components of the human epidermal cornified cell envelope. The Journal of biological chemistry. 1995;270(30):17702-11.

33. Robinson NA, Lapic S, Welter JF, Eckert RL. S100A11, S100A10, annexin I, desmosomal proteins, small proline-rich proteins, plasminogen activator inhibitor-2, and involucrin are components of the cornified envelope of cultured human epidermal keratinocytes. The Journal of biological chemistry. 1997;272(18):12035-46.

34. Russell LJ, DiGiovanna JJ, Rogers GR, Steinert PM, Hashem N, Compton JG, et al. Mutations in the gene for transglutaminase 1 in autosomal recessive lamellar ichthyosis. Nature genetics. 1995;9(3):279-83.

35. Huber M, Rettler I, Bernasconi K, Frenk E, Lavrijsen SP, Ponec M, et al. Mutations of keratinocyte transglutaminase in lamellar ichthyosis. Science (New York, NY). 1995;267(5197):525-8.

36. Matsuki M, Yamashita F, Ishida-Yamamoto A, Yamada K, Kinoshita C, Fushiki S, et al. Defective stratum corneum and early neonatal death in mice lacking the gene for transglutaminase 1 (keratinocyte transglutaminase). Proceedings of the National Academy of Sciences of the United States of America. 1998;95(3):1044-9.

37. Maestrini E, Monaco AP, McGrath JA, Ishida-Yamamoto A, Camisa C, Hovnanian A, et al. A molecular defect in loricrin, the major component of the cornified cell envelope, underlies Vohwinkel's syndrome. Nature genetics. 1996;13(1):70-7.

38. Akiyama M. Corneocyte lipid envelope (CLE), the key structure for skin barrier function and ichthyosis pathogenesis. Journal of dermatological science. 2017;88(1):3-9.

39. Rabionet M, Gorgas K, Sandhoff R. Ceramide synthesis in the epidermis. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids. 2014;1841(3):422-34.

40. Grall A, Guaguere E, Planchais S, Grond S, Bourrat E, Hausser I, et al. PNPLA1 mutations cause autosomal recessive congenital ichthyosis in golden retriever dogs and humans. Nature genetics. 2012;44(2):140-7.

41. Ohno Y, Kamiyama N, Nakamichi S, Kihara A. PNPLA1 is a transacylase essential for the generation of the skin barrier lipid omega-O-acylceramide. Nature communications. 2017;8:14610.

42. Matsumoto M, Zhou Y, Matsuo S, Nakanishi H, Hirose K, Oura H, et al. Targeted deletion of the murine corneodesmosin gene delineates its essential role in skin and hair physiology. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(18):6720-4.

43. Levy-Nissenbaum E, Betz RC, Frydman M, Simon M, Lahat H, Bakhan T, et al. Hypotrichosis simplex of the scalp is associated with nonsense mutations in CDSN encoding corneodesmosin. Nature genetics. 2003;34(2):151-3.

44. Descargues P, Deraison C, Bonnart C, Kreft M, Kishibe M, Ishida-Yamamoto A, et al. Spink5-deficient mice mimic Netherton syndrome through degradation of desmoglein 1 by epidermal protease hyperactivity. Nature genetics. 2005;37(1):56-65.

45. Caubet C, Jonca N, Brattsand M, Guerrin M, Bernard D, Schmidt R, et al. Degradation of corneodesmosome proteins by two serine proteases of the kallikrein family, SCTE/KLK5/hK5 and SCCE/KLK7/hK7. The Journal of investigative dermatology. 2004;122(5):1235-44.

46. Chavanas S, Bodemer C, Rochat A, Hamel-Teillac D, Ali M, Irvine AD, et al. Mutations in SPINK5, encoding a serine protease inhibitor, cause Netherton syndrome. Nature genetics. 2000;25(2):141-2.

47. Bitoun E, Chavanas S, Irvine AD, Lonie L, Bodemer C, Paradisi M, et al. Netherton syndrome: disease expression and spectrum of SPINK5 mutations in 21 families. The Journal of investigative dermatology. 2002;118(2):352-61.

48. Has C, Bruckner-Tuderman L. The genetics of skin fragility. Annual review of genomics and human genetics. 2014;15:245-68.

49. Turcan I, Jonkman MF. Blistering disease: insight from the hemidesmosome and other components of the dermal-epidermal junction. Cell and tissue research. 2015;360(3):545-69.

50. Boeira VL, Souza ES, Rocha Bde O, Oliveira PD, Oliveira Mde F, Rego VR, et al. Inherited epidermolysis bullosa: clinical and therapeutic aspects. Anais brasileiros de dermatologia. 2013;88(2):185-98.

51. Fine JD, Bruckner-Tuderman L, Eady RA, Bauer EA, Bauer JW, Has C, et al. Inherited epidermolysis bullosa: updated recommendations on diagnosis and classification. Journal of the American Academy of Dermatology. 2014;70(6):1103-26.

52. Terabayashi T, Hanada K. Genome instability syndromes caused by impaired DNA repair and aberrant DNA damage responses. Cell biology and toxicology. 2018;34(5):337-50.

53. Moriwaki S. Human DNA repair disorders in dermatology: A historical perspective, current concepts and new insight. Journal of dermatological science. 2016;81(2):77-84.

54. Lehmann AR. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie. 2003;85(11):1101-11.

55. Nichols AF, Ong P, Linn S. Mutations specific to the xeroderma pigmentosum group E Ddb- phenotype. The Journal of biological chemistry. 1996;271(40):24317-20.

56. Itoh T, Nichols A, Linn S. Abnormal regulation of DDB2 gene expression in xeroderma pigmentosum group E strains. Oncogene. 2001;20(48):7041-50.

57. Takahashi Y, Moriwaki S, Sugiyama Y, Endo Y, Yamazaki K, Mori T, et al. Decreased gene expression responsible for post-ultraviolet DNA repair synthesis in aging: a possible mechanism of age-related reduction in DNA repair capacity. The Journal of investigative dermatology. 2005;124(2):435-42.

58. Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, Yuasa M, et al. The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta. Nature. 1999;399(6737):700-4.

59. Lehmann AR, Kirk-Bell S, Arlett CF, Paterson MC, Lohman PH, de Weerd-Kastelein EA, et al. Xeroderma pigmentosum cells with normal levels of excision repair have a defect in DNA synthesis after UV-irradiation. Proceedings of the National Academy of Sciences of the United States of America. 1975;72(1):219-23.

60. Dessinioti C, Stratigos AJ, Rigopoulos D, Katsambas AD. A review of genetic disorders of hypopigmentation: lessons learned from the biology of melanocytes. Experimental dermatology. 2009;18(9):741-9.

61. Kaelin CB, Barsh GS. Genetics of pigmentation in dogs and cats. Annual review of animal biosciences. 2013;1:125-56.

62. Kawakami A, Fisher DE. The master role of microphthalmia-associated transcription factor in melanocyte and melanoma biology. Laboratory investigation; a journal of technical methods and pathology. 2017;97(6):649-56.

63. Tassabehji M, Newton VE, Read AP. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (MITF) gene. Nature genetics. 1994;8(3):251-5.

64. Tietz W. A syndrome of deaf-mutism associated with albinism showing dominant autosomal inheritance. American journal of human genetics. 1963;15:259-64.

65. George A, Zand DJ, Hufnagel RB, Sharma R, Sergeev YV, Legare JM, et al. Biallelic Mutations in MITF Cause Coloboma, Osteopetrosis, Microphthalmia, Macrocephaly, Albinism, and Deafness. American journal of human genetics. 2016;99(6):1388-94.

66. Hauswirth R, Haase B, Blatter M, Brooks SA, Burger D, Drogemuller C, et al. Mutations in MITF and PAX3 cause "splashed white" and other white spotting phenotypes in horses. PLoS genetics. 2012;8(4):e1002653.

67. Gronskov K, Ek J, Brondum-Nielsen K. Oculocutaneous albinism. Orphanet journal of rare diseases. 2007;2:43.

68. Wolf Horrell EM, Boulanger MC, D'Orazio JA. Melanocortin 1 Receptor: Structure, Function, and Regulation. Frontiers in genetics. 2016;7:95.

69. Newton JM, Wilkie AL, He L, Jordan SA, Metallinos DL, Holmes NG, et al. Melanocortin 1 receptor variation in the domestic dog. Mammalian genome : official journal of the International Mammalian Genome Society. 2000;11(1):24-30.

70. Marks MS, Seabra MC. The melanosome: membrane dynamics in black and white. Nature reviews Molecular cell biology. 2001;2(10):738-48.

71. Raposo G, Marks MS. Melanosomes--dark organelles enlighten endosomal membrane transport. Nature reviews Molecular cell biology. 2007;8(10):786-97.

72. Clark LA, Wahl JM, Rees CA, Murphy KE. Retrotransposon insertion in SILV is responsible for merle patterning of the domestic dog. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(5):1376-81.

73. Oh J, Bailin T, Fukai K, Feng GH, Ho L, Mao JI, et al. Positional cloning of a gene for Hermansky-Pudlak syndrome, a disorder of cytoplasmic organelles. Nature genetics. 1996;14(3):300-6.

74. Sckolnick M, Krementsova EB, Warshaw DM, Trybus KM. More than just a cargo adapter, melanophilin prolongs and slows processive runs of myosin Va. The Journal of biological chemistry. 2013;288(41):29313-22.

75. Tadokoro R, Takahashi Y. Intercellular transfer of organelles during body pigmentation. Current opinion in genetics & development. 2017;45:132-8.

76. Tadokoro R, Murai H, Sakai KI, Okui T, Yokota Y, Takahashi Y. Melanosome transfer to keratinocyte in the chicken embryonic skin is mediated by vesicle release associated with Rho-regulated membrane blebbing. Scientific reports. 2016;6:38277.

77. Scott G, Leopardi S, Parker L, Babiarz L, Seiberg M, Han R. The proteinase-activated receptor-2 mediates phagocytosis in a Rho-dependent manner in human keratinocytes. The Journal of investigative dermatology. 2003;121(3):529-41.

78. Mikkola ML, Millar SE. The mammary bud as a skin appendage: unique and shared aspects of development. Journal of mammary gland biology and neoplasia. 2006;11(3-4):187-203.

79. Mikkola ML. Genetic basis of skin appendage development. Seminars in cell & developmental biology. 2007;18(2):225-36.

80. Casal ML, Lewis JR, Mauldin EA, Tardivel A, Ingold K, Favre M, et al. Significant correction of disease after postnatal administration of recombinant ectodysplasin A in canine X-linked ectodermal dysplasia. American journal of human genetics. 2007;81(5):1050-6.

81. Mikkola ML. TNF superfamily in skin appendage development. Cytokine & growth factor reviews. 2008;19(3-4):219-30.

82. Shirokova V, Jussila M, Hytonen MK, Perala N, Drogemuller C, Leeb T, et al. Expression of Foxi3 is regulated by ectodysplasin in skin appendage placodes. Developmental dynamics : an official publication of the American Association of Anatomists. 2013;242(6):593-603.

83. Drogemuller C, Distl O, Leeb T. Partial deletion of the bovine ED1 gene causes anhidrotic ectodermal dysplasia in cattle. Genome research. 2001;11(10):1699-705.

84. Fuchs E. Scratching the surface of skin development. Nature. 2007;445(7130):834-42.

85. Rendl M, Lewis L, Fuchs E. Molecular dissection of mesenchymal-epithelial interactions in the hair follicle. PLoS biology. 2005;3(11):e331.

86. Sennett R, Rendl M. Mesenchymal-epithelial interactions during hair follicle morphogenesis and cycling. Seminars in cell & developmental biology. 2012;23(8):917-27.

87. Duverger O, Morasso MI. To grow or not to grow: hair morphogenesis and human genetic hair disorders. Seminars in cell & developmental biology. 2014;25-26:22-33.

88. Winter H, Rogers MA, Langbein L, Stevens HP, Leigh IM, Labreze C, et al. Mutations in the hair cortex keratin hHb6 cause the inherited hair disease monilethrix. Nature genetics. 1997;16(4):372-4.

89. Winter H, Rogers MA, Gebhardt M, Wollina U, Boxall L, Chitayat D, et al. A new mutation in the type II hair cortex keratin hHb1 involved in the inherited hair disorder monilethrix. Human genetics. 1997;101(2):165-9.

90. Winter H, Labreze C, Chapalain V, Surleve-Bazeille JE, Mercier M, Rogers MA, et al. A variable monilethrix phenotype associated with a novel mutation, Glu402Lys, in the helix termination motif of the type II hair keratin hHb1. The Journal of investigative dermatology. 1998;111(1):169-72.

91. Schweizer J, Langbein L, Rogers MA, Winter H. Hair follicle-specific keratins and their diseases. Experimental cell research. 2007;313(10):2010-20.

92. Welle MM, Wiener DJ. The Hair Follicle: A Comparative Review of Canine Hair Follicle Anatomy and Physiology. Toxicologic pathology. 2016;44(4):564-74.

93. Brouillard P, Vikkula M. Genetic causes of vascular malformations. Human molecular genetics. 2007;16 Spec No. 2:R140-9.

94. Brouillard P, Vikkula M. Vascular malformations: localized defects in vascular morphogenesis. Clinical genetics. 2003;63(5):340-51.

95. Srinivasan RS, Dillard ME, Lagutin OV, Lin FJ, Tsai S, Tsai MJ, et al. Lineage tracing demonstrates the venous origin of the mammalian lymphatic vasculature. Genes & development. 2007;21(19):2422-32.

96. Martinez-Corral I, Ulvmar MH, Stanczuk L, Tatin F, Kizhatil K, John SW, et al. Nonvenous origin of dermal lymphatic vasculature. Circulation research. 2015;116(10):1649-54.

97. Adams RH, Alitalo K. Molecular regulation of angiogenesis and lymphangiogenesis. Nature reviews Molecular cell biology. 2007;8(6):464-78.

98. Govani FS, Shovlin CL. Hereditary haemorrhagic telangiectasia: a clinical and scientific review. European journal of human genetics : EJHG. 2009;17(7):860-71.

99. Wooderchak-Donahue WL, McDonald J, O'Fallon B, Upton PD, Li W, Roman BL, et al. BMP9 mutations cause a vascular-anomaly syndrome with phenotypic overlap with hereditary hemorrhagic telangiectasia. American journal of human genetics. 2013;93(3):530-7.

100. McDonald J, Bayrak-Toydemir P, Pyeritz RE. Hereditary hemorrhagic telangiectasia: an overview of diagnosis, management, and pathogenesis. Genetics in medicine : official journal of the American College of Medical Genetics. 2011;13(7):607-16.

101. Vanakker O, Callewaert B, Malfait F, Coucke P. The Genetics of Soft Connective Tissue Disorders. Annual review of genomics and human genetics. 2015;16:229-55.

102. Nystrom A, Bruckner-Tuderman L. Matrix molecules and skin biology. Seminars in cell & developmental biology. 2018.

103. Morava E, Wopereis S, Coucke P, Gillessen-Kaesbach G, Voit T, Smeitink J, et al. Defective protein glycosylation in patients with cutis laxa syndrome. European journal of human genetics : EJHG. 2005;13(4):414-21.

104. Malfait F, Francomano C, Byers P, Belmont J, Berglund B, Black J, et al. The 2017 international classification of the Ehlers-Danlos syndromes. American journal of medical genetics Part C, Seminars in medical genetics. 2017;175(1):8-26.

105. Beighton P. Ehlers-Danlos syndrome. Annals of the rheumatic diseases. 1970;29(3):332-3.

106. Myllyharju J, Kivirikko KI. Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends in genetics : TIG. 2004;20(1):33-43.

107. Kielty CM, E. GM. The Collagen Family: Structure, Assembly, and Organization in the Extracellular Matrix. In: M. RP, Beat S, editors. Connective Tissue and Its Heritable Disorders. 2nd ed. New York: Wiley-Liss; 2002.

108. Bateman JF, Boot-Handford RP, Lamande SR. Genetic diseases of connective tissues: cellular and extracellular effects of ECM mutations. Nature reviews Genetics. 2009;10(3):173-83.

109. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. Nature reviews Genetics. 2013;14(5):307-20.

110. Lim YH, Moscato Z, Choate KA. Mosaicism in Cutaneous Disorders. Annual review of genetics. 2017;51:123-41.

111. Cohen MM, Jr. Proteus syndrome review: molecular, clinical, and pathologic features. Clinical genetics. 2014;85(2):111-9.

112. Lindhurst MJ, Sapp JC, Teer JK, Johnston JJ, Finn EM, Peters K, et al. A mosaic activating mutation in AKT1 associated with the Proteus syndrome. The New England journal of medicine. 2011;365(7):611-9.

113. Manning BD, Toker A. AKT/PKB Signaling: Navigating the Network. Cell. 2017;169(3):381-405.

114. Happle R. The categories of cutaneous mosaicism: A proposed classification. American journal of medical genetics Part A. 2016;170a(2):452-9.

115. Vreeburg M, van Steensel MA. Genodermatoses caused by genetic mosaicism. European journal of pediatrics. 2012;171(12):1725-35.

116. Jaju PD, Ransohoff KJ, Tang JY, Sarin KY. Familial skin cancer syndromes: Increased risk of nonmelanotic skin cancers and extracutaneous tumors. Journal of the American Academy of Dermatology. 2016;74(3):437-51; quiz 52-4.

117. Zhang X, Luo S, Wu J, Zhang L, Wang WH, Degan S, et al. KIND1 Loss Sensitizes Keratinocytes to UV-Induced Inflammatory Response and DNA Damage. The Journal of investigative dermatology. 2017;137(2):475-83.

118. Armanios M, Chen JL, Chang YP, Brodsky RA, Hawkins A, Griffin CA, et al. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(44):15960-4.

119. Dokal I. Dyskeratosis congenita in all its forms. British journal of haematology. 2000;110(4):768-79.

120. Kirwan M, Dokal I. Dyskeratosis congenita: a genetic disorder of many faces. Clinical genetics. 2008;73(2):103-12.

121. Hartwig FP, Collares T. Telomere dysfunction and tumor suppression responses in dyskeratosis congenita: balancing cancer and tissue renewal impairment. Ageing research reviews. 2013;12(2):642-52.

122. Collins FS. Positional cloning moves from perditional to traditional. Nature genetics. 1995;9(4):347-50.

123. Teare MD, Santibanez Koref MF. Linkage analysis and the study of Mendelian disease in the era of whole exome and genome sequencing. Briefings in functional genomics. 2014;13(5):378-83.

124. Burton PR, Tobin MD, Hopper JL. Key concepts in genetic epidemiology. Lancet (London, England). 2005;366(9489):941-51.

125. Ott J, Wang J, Leal SM. Genetic linkage analysis in the age of whole-genome sequencing. Nature reviews Genetics. 2015;16(5):275-84.

126. Vink JM, Boomsma DI. Gene finding strategies. Biological psychology. 2002;61(1-2):53-71.

127. Vahidnezhad H, Youssefian L, Jazayeri A, Uitto J. Research Techniques Made Simple: Genome-Wide Homozygosity/Autozygosity Mapping Is a Powerful Tool for Identifying Candidate Genes in Autosomal Recessive Genetic Diseases. The Journal of investigative dermatology. 2018;138(9):1893-900.

128. Teo YY. Common statistical issues in genome-wide association studies: a review on power, data quality control, genotype calling and population structure. Current opinion in lipidology. 2008;19(2):133-43.

129. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nature reviews Genetics. 2014;15(5):335-46.

130. de los Campos G, Vazquez AI, Hsu S, Lello L. Complex-Trait Prediction in the Era of Big Data. Trends in Genetics. 2018.

131. Karlsson EK, Lindblad-Toh K. Leader of the pack: gene mapping in dogs and other model organisms. Nature reviews Genetics. 2008;9(9):713-25.

132. Gao X, Becker LC, Becker DM, Starmer JD, Province MA. Avoiding the high Bonferroni penalty in genome-wide association studies. Genetic epidemiology. 2010;34(1):100-5.

133. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(16):9440-5.

134. Han B, Kang HM, Eskin E. Rapid and accurate multiple testing correction and power estimation for millions of correlated markers. PLoS genetics. 2009;5(4):e1000456.

135. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. Lancet (London, England). 2003;361(9357):598-604.

136. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature genetics. 2006;38(2):203-8.

137. Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. Nature reviews Genetics. 2010;11(7):459-63.

138. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America. 1977;74(12):5463-7.

139. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. Cell. 2013;155(1):27-38.

140. Djakow J, Kramna L, Dusatkova L, Uhlik J, Pursiheimo JP, Svobodova T, et al. An effective combination of sanger and next generation sequencing in diagnostics of primary ciliary dyskinesia. Pediatric pulmonology. 2016;51(5):498-509.

141. McGinn S, Gut IG. DNA sequencing - spanning the generations. New biotechnology. 2013;30(4):366-72.

142. Takeichi T, Nanda A, Liu L, Salam A, Campbell P, Fong K, et al. Impact of next generation sequencing on diagnostics in a genetic skin disease clinic. Experimental dermatology. 2013;22(12):825-31.

143. Erguner B, Ustek D, Sagiroglu MS. Performance comparison of Next Generation sequencing platforms. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference. 2015;2015:6453-6.

144. Maruthappu T, Scott CA, Kelsell DP. Discovery in genetic skin disease: the impact of high throughput genetic technologies. Genes. 2014;5(3):615-34.

145. Jagannathan V, Bannoehr J, Plattet P, Hauswirth R, Drogemuller C, Drogemuller M, et al. A mutation in the SUV39H2 gene in Labrador Retrievers with hereditary nasal parakeratosis (HNPK) provides insights into the epigenetics of keratinocyte differentiation. PLoS genetics. 2013;9(10):e1003848.

146. Ovaere P, Lippens S, Vandenabeele P, Declercq W. The emerging roles of serine protease cascades in the epidermis. Trends in biochemical sciences. 2009;34(9):453-63.

147. Volkel P, Angrand PO. The control of histone lysine methylation in epigenetic regulation. Biochimie. 2007;89(1):1-20.

148. Paller AS, van Steensel MA, Rodriguez-Martin M, Sorrell J, Heath C, Crumrine D, et al. Pathogenesis-based therapy reverses cutaneous abnormalities in an inherited disorder of distal cholesterol metabolism. The Journal of investigative dermatology. 2011;131(11):2242-8.

149. Oberhofer A, Spieler P, Rosenfeld Y, Stepp WL, Cleetus A, Hume AN, et al. Myosin Va's adaptor protein melanophilin enforces track selection on the microtubule and actin networks in vitro. Proceedings of the National Academy of Sciences of the United States of America. 2017;114(24):E4714-e23.

150. Bellono NW, Oancea EV. Ion transport in pigmentation. Archives of biochemistry and biophysics. 2014;563:35-41.

151. Kadler KE, Hill A, Canty-Laird EG. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. Curr Opin Cell Biol. 2008;20(5):495-501.

152. Lefebvre S, Mikkola ML. Ectodysplasin research--where to next? Seminars in immunology. 2014;26(3):220-8.

153. Bauer A, Nimmo J, Newman R, Brunner M, Welle MM, Jagannathan V, et al. A splice site variant in the SUV39H2 gene in Greyhounds with nasal parakeratosis. Animal genetics. 2018;49(2):137-40.

154. Bauer A, De Lucia M, Jagannathan V, Mezzalira G, Casal ML, Welle MM, et al. A Large Deletion in the NSDHL Gene in Labrador Retrievers with a Congenital Cornification Disorder. G3 (Bethesda, Md). 2017;7(9):3115-21.

155. Spycher M, Bauer A, Jagannathan V, Frizzi M, De Lucia M, Leeb T. A frameshift variant in the COL5A1 gene in a cat with Ehlers-Danlos syndrome. Animal genetics. 2018. doi: 10.1111/age.12727.

156. Caduff M, Bauer A, Jagannathan V, Leeb T. A single base deletion in the SLC45A2 gene in a Bullmastiff with oculocutaneous albinism. Animal genetics. 2017;48(5):619-21.

157. Bauer A, Waluk DP, Galichet A, Timm K, Jagannathan V, Sayar BS, et al. A de novo variant in the ASPRV1 gene in a dog with ichthyosis. PLoS genetics. 2017;13(3):e1006651.

158. Bauer A, Kehl A, Jagannathan V, Leeb T. A novel MLPH variant in dogs with coat colour dilution. Animal genetics. 2018;49(1):94-7.

159. Bauer A, Jagannathan V, Hogler S, Richter B, McEwan NA, Thomas A, et al. MKLN1 splicing defect in dogs with lethal acrodermatitis. PLoS genetics. 2018;14(3):e1007264.

160. Bauer A, Hiemesch T, Jagannathan V, Neuditschko M, Bachmann I, Rieder S, et al. A Nonsense Variant in the ST14 Gene in Akhal-Teke Horses with Naked Foal Syndrome. G3 (Bethesda, Md). 2017;7(4):1315-21.

161. Hadji Rasouliha S, Bauer A, Dettwiler M, Welle MM, Leeb T. A frameshift variant in the EDA gene in Dachshunds with X-linked hypohidrotic ectodermal dysplasia. Animal genetics. 2018. doi: 10.1111/age.12729.

162. Adams JC, Seed B, Lawler J. Muskelin, a novel intracellular mediator of cell adhesive and cytoskeletal responses to thrombospondin-1. The EMBO journal. 1998;17(17):4964-74.

163. Francis O, Han F, Adams JC. Molecular phylogeny of a RING E3 ubiquitin ligase, conserved in eukaryotic cells and dominated by homologous components, the muskelin/RanBPM/CTLH complex. PloS one. 2013;8(10):e75217.

164. Ledee DR, Gao CY, Seth R, Fariss RN, Tripathi BK, Zelenka PS. A specific interaction between muskelin and the cyclin-dependent kinase 5 activator p39 promotes peripheral localization of muskelin. The Journal of biological chemistry. 2005;280(22):21376-83.

165. Hasegawa H, Katoh H, Fujita H, Mori K, Negishi M. Receptor isoform-specific interaction of prostaglandin EP3 receptor with muskelin. Biochemical and biophysical research communications. 2000;276(1):350-4.

166. Gueron G, Giudice J, Valacco P, Paez A, Elguero B, Toscani M, et al. Heme-oxygenase-1 implications in cell morphology and the adhesive behavior of prostate cancer cells. Oncotarget. 2014;5(12):4087-102.

167. Debenedittis P, Harmelink C, Chen Y, Wang Q, Jiao K. Characterization of the novel interaction between muskelin and TBX20, a critical cardiogenic transcription factor. Biochemical and biophysical research communications. 2011;409(2):338-43.

168. Valiyaveettil M, Bentley AA, Gursahaney P, Hussien R, Chakravarti R, Kureishy N, et al. Novel role of the muskelin-RanBP9 complex as a nucleocytoplasmic mediator of cell morphology regulation. The Journal of cell biology. 2008;182(4):727-39.

169. Kobayashi N, Yang J, Ueda A, Suzuki T, Tomaru K, Takeno M, et al. RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8alpha and ARMC8beta are components of the CTLH complex. Gene. 2007;396(2):236-47.

170. Heisler FF, Pechmann Y, Wieser I, Altmeppen HC, Veenendaal L, Muhia M, et al. Muskelin Coordinates PrP(C) Lysosome versus Exosome Targeting and Impacts Prion Disease Progression. Neuron. 2018;99(6):1155-69.e9.

171. Heisler FF, Loebrich S, Pechmann Y, Maier N, Zivkovic AR, Tokito M, et al. Muskelin regulates actin filament- and microtubule-based GABA(A) receptor transport in neurons. Neuron. 2011;70(1):66-81.

172. Konig A, Happle R, Fink-Puches R, Soyer HP, Bornholdt D, Engel H, et al. A novel missense mutation of NSDHL in an unusual case of CHILD syndrome showing bilateral, almost symmetric involvement. Journal of the American Academy of Dermatology. 2002;46(4):594-6.

173. Happle R, Mittag H, Kuster W. The CHILD nevus: a distinct skin disorder. Dermatology (Basel, Switzerland). 1995;191(3):210-6.

174. Tantcheva-Poor I, Oji V, Has C. A multistep approach to the diagnosis of rare genodermatoses. Journal der Deutschen Dermatologischen Gesellschaft = Journal of the German Society of Dermatology : JDDG. 2016;14(10):969-86.

175. Hytonen MK, Lohi H. Canine models of human rare disorders. Rare diseases (Austin, Tex). 2016;4(1):e1241362.

176. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, et al. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature. 2005;438(7069):803-19.

177. Sutter NB, Ostrander EA. Dog star rising: the canine genetic system. Nature reviews Genetics. 2004;5(12):900-10.

178. Patterson DF. Companion animal medicine in the age of medical genetics. Journal of veterinary internal medicine. 2000;14(1):1-9.

179. Tsai KL, Clark LA, Murphy KE. Understanding hereditary diseases using the dog and human as companion model systems. Mammalian genome : official journal of the International Mammalian Genome Society. 2007;18(6-7):444-51.

180. Sutter NB, Eberle MA, Parker HG, Pullar BJ, Kirkness EF, Kruglyak L, et al. Extensive and breed-specific linkage disequilibrium in Canis familiaris. Genome research. 2004;14(12):2388-96.

181. Hoeppner MP, Lundquist A, Pirun M, Meadows JR, Zamani N, Johnson J, et al. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. PloS one. 2014;9(3):e91172.

182. Walters-Conte KB, Johnson DL, Allard MW, Pecon-Slattery J. Carnivore-specific SINEs (Can-SINEs): distribution, evolution, and genomic impact. The Journal of heredity. 2011;102 Suppl 1:S2-10.

183. Di Rienzo A. Population genetics models of common diseases. Current opinion in genetics & development. 2006;16(6):630-6.

184. Marian AJ. Causality in genetics: the gradient of genetic effects and back to Koch's postulates of causality. Circulation research. 2014;114(2):e18-21.

Declaration of Originality

Last name, first name: Bauer Anina Estrella

Matriculation number: 10-102-119

I hereby declare that this thesis represents my original work and that I have used no other sources except as noted by citations.

All data, tables, figures and text citations which have been reproduced from any other source, including the internet, have been explicitly acknowledged as such.

I am aware that in case of non-compliance, the Senate is entitled to withdraw the doctorate degree awarded to me on the basis of the present thesis, in accordance with the "Statut der Universität Bern (Universitätsstatut; UniSt)", Art. 69, of 7 June 2011.

Place, date

Bern, 20 December 2018

Signature